

Seedless in Seattle?

When negotiators gather in Seattle this weekend for the ministerial meeting of the World Trade Organization, the future of genetically modified crops will be high on their agenda.

The deliberations of the World Trade Organization (WTO), in the context of the global economy, carry far more weight than the legislative bodies of many sovereign governments. The path that the organization chooses at Seattle will therefore play a major part in determining the international viability of transgenic technology. Government representatives and others attending the WTO's third biennial ministerial meeting will be under pressure to handle the issue with extreme care, aware that differences between the United States and Europe could easily spark a damaging trade war. But the WTO meeting will feature divisions not just between the United States and Europe, but between rich and poor nations.

There are two areas of negotiation at Seattle where the interests of rich and poor nations are set to clash. The first concerns a proposal, originally put forward by Canada and Japan, to establish a 'working party' at the WTO expressly to deal with biotechnology issues. The casual observer might ask, who could possibly object? We all admire work and like parties, and deferring a thorny problem for further study never did anyone any harm.

In this case, however, developing countries are viewing the working party with considerable suspicion. Because the WTO has become so powerful, some fear that its formal involvement in the question of agricultural biotechnology could undermine other international agreements — such as any biosafety protocol that may one day be negotiated under the Convention on Biological Diversity — that are potentially more sensitive to environmental concerns.

The biodiversity convention carries at least the potential of helping developing countries to conserve and exploit their substantial biological resources. They are understandably concerned that WTO involvement will lead to environmental considerations being swamped by those of trade. The remit of any WTO group dealing with agricultural biotechnology must therefore be carefully circum-

scribed, at the very least. Free trade, whatever certain economists, politicians and professional trade negotiators at the WTO may believe, is not more important than the future of life on our planet.

The second source of North–South friction concerns the future of the Trade Related Intellectual Property Rights (TRIPs) agreement, as it applies to the patenting of plants and animals. TRIPs tries to globalize the type of intellectual property rights regime that protects innovation in the United States and Europe. The fine print of the agreement gives signatories until the end of this year to review a clause allowing a partial exemption for plants and animals.

At Seattle, the United States will be trying to maintain the status quo, while ensuring that the exemption is implemented in a way that allows substantive protection of genetic modifications of both types of organisms. But many developing countries want to expand the clause to bar the patenting of all forms of life. Next year, the whole TRIPs agreement is up for review, and developing countries will seek to revise what they view as its exploitative aspects.

The genes and methods behind agricultural biotechnology are already caught in a web of patent protection that affects the ability of researchers in developing countries to deliver the technology to poor farmers (see Briefing, pages 341–345). The Seattle meeting will be besieged by tens of thousands of demonstrators protesting against the entire negotiation as an affront to the world's poor, and heavily lobbied by agribusiness interests, pushing for open markets and tough controls of intellectual property rights. The best that can be hoped for in Seattle is a compromise deal, establishing a trading regime that will allow genetically modified food seed to enter markets at a pace acceptable to consumers, and plant breeders in developing countries to work with transgenic crop technology at an affordable cost. Even this modest goal, however, would be a substantial achievement in the current, highly charged atmosphere. ■

A bioethics dilemma for Germany

Debates on the ethics of biomedical research must find a way to combine breadth of vision with a practical timescale.

At first glance, the backtracking of Germany's Social Democrat Party on a pre-election promise to support an inquiry into the social impact of modern genetics and biomedicine (see page 331) seems perverse. Similar inquiries by comparable bodies in other countries have been successful in using focused debate, backed by careful research, to forge a workable consensus on such contentious topics. And the proposed procedure — a so-called Enquete Commission — was designed specifically for such a task.

But the experience of such inquiries in Germany has been mixed. Some, such as one into genetic engineering in the mid-1980s, have been successful, while others — for example, into technology assessment — have not. As the latter showed, the specific form of debate encouraged, however democratic in principle and practice, can fail to mesh with a practical timescale. Where issues are pressing, or where — as in the case of the implementation of the Council of Europe's Declaration on Human Rights and Biomedicine — other European states are setting the pace, an open-ended debate intended to

achieve a conclusion through reasoned discussion alone can be an unrealistic strategy.

To that extent, the Social Democrats, their members in the government, and other political parties, are right to explore more practical alternatives. One is to create a national bioethics advisory council with broad responsibilities to monitor specific bioethical issues as these arise, for example through the activities of targeted working groups. Another would be to explore the 'consensus conference' model, allowing non-experts to form judgements based on a critical review of the evidence.

At the same time, however, it would be short-sighted if any alternative to the commission was to be so narrowly focused as to lose sight of the broader perspective. In that sense, the Greens, despite their overly romantic, anti-technology stance, are right: whatever its faults, an Enquete Commission allows for a range of views to be discussed with a breadth and seriousness appropriate to the significance of the topics under debate. This must not be lost. ■





Meteor forecast
The Leonid meteor storm's peak was predicted precisely
p333



Reactor reaction
Brookhaven's High Flux Beam Reactor is to close for good
p333



Big spenders
Howard Hughes Medical Institute names new team
p334



Launch failure
Further setback for Japan's space programme
p336

German bioethics inquiry 'could hold up essential rule changes'

Munich

In contrast to their pre-election promise, the Social Democrats (SPD) in the German parliament have declared their opposition to an all-parliamentary commission of inquiry on human rights and bioethics.

They argue that such a '*commission d'enquête*', on whose creation parliament is due to vote next week, could delay the introduction of rules and legislation in the fields of biomedicine by producing lengthy debate on principles that have already been agreed.

Their position is strongly supported by the Christian Democrat opposition. But the Greens — the SPD's junior coalition partner — complain that refusing to set up the promised commission would be shirking a democratic responsibility.

"Advances in pre-implantation diagnostics and reproductive medicine are likely to shatter our traditional ideas of human nature," says Monika Knoche, the Green parliamentary group's expert on medical ethics. "A democratic parliament has an obligation to take up such issues."

Parliamentary *commissions d'enquête*, intended to provide broad advice to politicians, are set up only on issues with major impact on society, such as climate change, economic globalization or the consequences of demographic changes. They comprise members of all political parties, as well as representatives of relevant social groups.

But their work can be lengthy and time-consuming, and sometimes a final report, as occurred with that on technology assessment (*Technikfolgenabschätzung*), cannot be reached within a four-year legislation period. The wide range of topics discussed, and the conflicting perspectives of groups and individuals involved, makes compromise difficult to achieve.

Some argue that this makes *commissions d'enquête* a useful way of postponing political action, such as legislation. Furthermore, given the country's Nazi past, medical ethics issues are already highly sensitive in Germany, and its legislation is the strictest in Europe (*Nature* **389**, 660; 1997 & **384**, 5; 1996).

Germany is still hesitating, for example, over whether to sign and ratify the Council of

Europe's Convention on Human Rights and Biomedicine. It fears that its own laws, although not directly in conflict with the convention, could eventually be watered down by its less restrictive provisions.

For many Germans, the convention's main bone of contention is a clause allowing research on 'legally incapacitated' persons where there is no alternative means of doing the research.

Such research is forbidden in Germany under the Nuremberg codex of physicians. However, it is being carried out in what experts describe as a legal 'grey area' — for example, clinical research on oxygen supply during birth complications, with the consent of the babies' parents alone.

Ironically, the SPD's refusal to set up a *commission d'enquête* on bioethics reflects the party's attempt to prevent similar grey

areas arising in other areas of biomedicine, such as germline therapy, genetic testing, reproduction medicine, organ transplantation or the cloning of stem cells.

"We should spend our time taking legislative initiatives where they are needed, rather than engaging in a cumbersome debate about principles," says Wolf-Michael Catenhusen, the SPD's parliamentary state secretary in the Ministry of Research.

Catenhusen favours signing the Council of Europe convention. The convention's supplementary agreements on organ transplantation, embryo research, human genetics and medical research are now being discussed by the council's bioethics steering committee. Germany is involved in the discussions, but Catenhusen is worried about what may come out of these.

So far, the convention, approved by the

'It's a G': the one-billionth nucleotide

Washington & Cambridge

Champagne corks were popping on both sides of the Atlantic on Tuesday (23 November) when participants in the Human Genome Project celebrated the successful sequencing of the one-billionth base pair of human DNA.

The landmark means that researchers are approximately one third of the way towards the full sequence of about three billion nucleotides, which is due to be completed next spring. According to officials at the Wellcome Trust in London, the one-billionth nucleotide, reached on 17 November, was a 'G' (guanine).

At a ceremony at the National Academy of Sciences in Washington, Donna Shalala, the US Secretary for Health and Human Services, handed certificates to representatives of the main US sequencing centres, each of which was linked by video to the award ceremony. Shalala paid tribute to "the brilliance, dedication and ingenuity of hundreds of scientists throughout the world. They've been doing this quietly,



Sulston: received award for work on genome.

but at an astonishing pace, and their work promises to fuel unprecedented scientific and medical advances."

Francis Collins, director of the National Human

Genome Research Institute, which is supporting the US part of the project with the energy department, praised participants for keeping costs down and quality up.

Britain's science minister, David Sainsbury, told a simultaneous celebration at the Sanger Centre outside Cambridge, where one third of the sequencing is being carried out, that the speed and skill with which the one-billionth pair had been reached was "a remarkable achievement". Earlier in the day, Sainsbury presented an award to John Sulston, the director of the Sanger Centre.

David Dickson & Colin Macilwain

► council in 1997, has been signed by 28 of the council's 40 member states. Five countries — Denmark, Greece, San Marino, Slovakia and Slovenia — have ratified it, and Spain will join them in January.

Germany and Britain are the only large European countries not to have signed. "I feel that Germany should not be the last to sign," says Catenhusen. "Otherwise we risk being excluded from the discussion process in Europe."

The Greens, at their last party conference, unanimously voted against the convention. They fear it could undermine



Catenhusen: 'sign European convention'.

German standards, bringing them down to the levels of some neighbouring countries such as Belgium, where, for example, there are no laws on embryo research.

But some members, such as Jens Reich, a bioinformaticist at the Max-Delbrück Centre for Molecular Medicine in Berlin and the Greens' 1994 candidate for federal president, argue more pragmatically that, despite its weaknesses, the convention should be signed as it sets minimal standards throughout Europe.

Ludger Honnefelder, director of the Institute of Science and Ethics at the University of Bonn, is a member of the German delegation in the Council of Europe's steering committee on bioethics. He firmly believes that Germany's accession to the convention would "significantly increase its international influence".

Despite the dispute about a *commission d'enquête*, there is general agreement in Germany on the need for continuous bioethical advice and monitoring. Catenhusen suggests that a national ethics commission, similar to Britain's Nuffield Council on Bioethics, should be established.

Werner Lensing, the Christian Democrat spokesman on bioethics in the parliamentary committee on research, agrees that ad hoc expert panels, which are used by the Nuffield council, would be effective. Just such an advisory panel, on the ethical implications of predictive genetic testing and reproductive medicine, was set up last week by the federal health ministry.

But Knoche insists that only a *commission d'enquête* would have the political weight appropriate to the scope of the issues in question: "Such a body is indispensable. And we would certainly not use it as a blocking tool."

Quirin Schiermeier

Israel urged to set up labs to carry out military research

Jerusalem

Government laboratories should be set up in Israel to carry out defence research, according to a report on revitalizing the country's military industries. Weapons development and production, it adds, should be carried out by private companies, rather than government corporations as at present.

The report was drawn up by a defence ministry committee headed by reserve general Moshe Peled. It recommends that Israel Aircraft Industries (IAI), Israel Military Industries, and Raphael (the Armament Development Authority), an auxiliary unit of the defence ministry, should be privatized.

But it adds that the research underpinning weapons development should be performed by a new system of government research laboratories.

As the report is classified, Peled is refusing to comment on it. But Yossi Snir, the committee's coordinator, says a summary that appeared in the Israeli daily newspaper *Ha'aretz* last week is "reliable".

In recent years, Israel's defence industries have been plagued by falling sales, labour disputes over forced redundancies, and a lack of technological innovation. According to *Ha'aretz*, one problem is the difficulty of attracting highly qualified young scientists.

"Since the 1990s [the trend] is for talented

people to leave government research laboratories and move to the flourishing private high-tech industry," confirms Zehev Tadmor, former president of the Technion, Israel's institute of technology.

There has been a mixed reaction to the proposal to create government defence laboratories, which do not currently exist in Israel. "I'm not sure that we have to separate out basic research," says Moshe Arens, former defence minister and former IAI deputy director-general. Privatizing the industries would be sufficient to make them more attractive to young scientists, he suggests.

But physicist and former minister of science and technology Yuval Ne'eman says the establishment of such laboratories would be a positive development, not only for defence industries but for Israeli science in general.

"Since nearly all research is performed at the universities, a tradition has become established according to which the amount of research performed in the country is tied to [its] number of students," Ne'eman complains. Government laboratories of the kind that exist in the United States and Europe could break the link between higher education and research budgets, he says.

The Peled report also calls for a reversal of the long decline in government investment in basic research.

Haim Watzman

Concern at cheap AIDS drug fears

Cape Town

Officials of a US foundation that has raised \$1 million to prevent paediatric AIDS in Africa are expressing concern at a statement from the health ministers of Southern African countries raising questions about the use of the anti-retroviral agent AZT and the cheaper alternative nevirapine.

The California-based Elizabeth Glaser Paediatric Trust has earmarked \$1 million for paediatric AIDS prevention in Africa using nevirapine, which is less expensive than AZT and simpler to administer. A single dose to mothers at the onset of labour, and a single dose to the baby in its first three days of life, cost less than \$4 per treatment.

But this initiative could suffer as the result of a joint statement issued by the health ministers of South Africa, Botswana, Zambia, Namibia, Mozambique, Swaziland, Lesotho, Zimbabwe, Malawi, Tanzania, Angola and Rwanda at a meeting in Johannesburg this month held to discuss a coordinated response on HIV/AIDS.

The ministers acknowledged that administering either drug can approximately halve the numbers of HIV-positive children born to mothers who are HIV positive.

But they expressed "grave concern over possible side effects as a result of their toxicity and the potential development of resistance to these compounds". They felt it was necessary to research the effects of "unnecessary exposure of children and mothers to these drugs".

This action is understood to have been heavily influenced by the South African government's position (see *Nature* 402, 225; 1999). In South Africa, 22 per cent of women attending antenatal clinics, and seven per cent of new babies, are HIV infected.

Both AZT and the oral form of nevirapine are registered with South Africa's Medicines Control Council. But it is understood that the suspension form of nevirapine, which is administered to infants, has not yet been submitted to the council for registration.

Michael Cherry

Praise greets accuracy of Leonid prediction

Washington

Last week's Leonid meteor storm, as well as being much more intense than last year's, marked the beginning of a new era in meteor forecasting. For the first time, researchers accurately predicted — to within minutes — when the two-day-long storm would peak.

The Leonids occur every November, when Earth's path crosses the tilted orbit of Comet Tempel–Tuttle. The display tends to be most dazzling at 33-year intervals, when the comet rounds the Sun and sheds a new layer of dust and grit to burn up in the Earth's atmosphere.

But forecasting the exact timing and intensity of any meteor shower (as measured by a standardized count known as the zenithal hourly rate, ZHR) is notoriously hard. Last year's Leonids, for example, surprised experts by peaking 16 hours earlier than expected.

Don Yeomans of the US space agency NASA's Jet Propulsion Laboratory, a veteran meteor prognosticator, wrote on his website last month that “like the weather, it is extremely difficult to predict the times and hourly rates of meteor showers”.

Yet David Asher, of Armagh Observatory in Northern Ireland, and Rob McNaught, of the Australian National University in Canberra, nailed this year's Leonids peak almost exactly. Their prediction of 2:08 Universal Time on 18 November was off by five minutes at the most.

Yeomans, basing his forecast on historic Leonids observations, missed by 20 minutes. Most researchers declined to hazard such a precise guess.

Asher and McNaught use computers to model dust emission from the comet, taking into account the rate and speed of ejection, particle size and other factors. They then add repeated gravitational tugs from Jupiter and other planets and watch how the dust disperses.

According to their model, the thin ribbons of dust created during each 33-year passage retain their coherence for centuries. Whereas scientists previously had assumed these independent dust trails follow the same path as the comet, Asher and McNaught found that the orbits are slightly offset.

Knowing their positions can give very accurate peak times, as well as the age of the source meteors in the shower. This year, the Earth passed through a cloud of dust shed from Tempel–Tuttle in 1899. Next year we'll hit the outskirts of the 1733 trail and the 1866 trail.

Asher and McNaught credit several Russian researchers, including E. D. Kondrat'eva and E. A. Reznikov, who pioneered this method of meteor forecasting in the



On cue: a Leonid meteor captured at 2:07 UT by the European Calar Alto observing team in Spain.

mid-1980s but went largely unnoticed by Western scientists.

Peter Brown of the University of Western Ontario, in London, Ontario, has obtained similar results with his own modelling of the Leonid stream. He also identified the 1899 cloud as the source of this year's shower.

The new modelling work has greatly advanced the art of meteor prediction, says Rainer Arlt of the Astrophysikalisches Institut Potsdam, who is compiling this year's Leonids observations for the International Meteor Organization. “All of these models are working, no matter what the details of each model are. We're really able to follow the dynamics over 1,000 years without problems.”

Predicting the meteor rate is a much tougher nut to crack, however. This year's ZHR was around 5,000, five to ten times the rate predicted by Asher and McNaught. Yeomans calls their forecast of the peak time “impressive”, but says the real proof will come in 2001 and 2002.

Yeomans and other experts say the Leonids should have pretty much died down by then, while Asher and McNaught are calling for intense storms with ZHRs of 15,000 or more. “That's going way out on a limb,” says Yeomans. “And if they get that right, I'll be the first to congratulate them on a major meteor-shower prediction breakthrough.”

Tony Reichhardt

Brookhaven leak reactor to close

Washington

The US Department of Energy announced last week that the High Flux Beam Reactor at Brookhaven National Laboratory on Long Island, New York, is to close for good. The 34-year-old research reactor is the main source of cold neutrons for researchers in the United States.

In announcing the reactor's closure, energy secretary Bill Richardson explained in a statement that “we need to focus our limited resources on productive research, rather than keeping the reactor in standby mode for an unknown length of time”.

But Brookhaven researchers reacted furiously to the sudden decision, which appeared to bypass a lengthy process that had been under way to assess the environmental impact of reopening the reactor. It has been closed since early 1997, when a small leak of tritium from the pool that stores its spent fuel was detected in the ground beneath it (see *Nature* 386, 3–4; 1997).

And although the eventual decision to close the reactor had been widely expected, neutron scientists were shocked that it was taken before the review process had been completed. “This has just become a circus,” says Denis McWhan, associate director for basic energy sciences at Brookhaven. “We all



Model pressure? ‘Supermodel’ Brinkley expressed opposition.

thought that we'd bought into a formal process, but the secretary just short-circuited it.” Bob Birgeneau, dean of science at the Massachusetts Institute of Technology, said he was “deeply

dismayed”. He predicted it would make US researchers “second- if not third-class citizens compared to Western Europe and Japan”.

Some researchers at Brookhaven linked the decision to a recent meeting between Richardson and two celebrated residents of Long Island, actor Alec Baldwin and ‘supermodel’ Christie Brinkley, who were representing a group opposing the reactor.

Richardson was in Turkey last week and unavailable for comment. But Ernie Moniz, the energy undersecretary, confirmed that such a meeting had taken place. “I believe he did meet with them. He got input from many different perspectives,” he said.

According to Moniz, however,



Off for good: the High Flux Beam Reactor.

Richardson's decision was based on the advice of programme managers, who included Patricia Dehmer, senior staff official in charge of basic energy sciences, and Martha Krebs, head of the energy department's Office of Science. They recommended against continuing to spend \$20 million a year to keep the facility on standby.

"He got the recommendation of a termination, and on that day he made the decision," says Moniz. He points out that Brookhaven will remain one of very few laboratories housing two major scientific facilities, the Relativistic Heavy Ion Collider and the National Synchrotron Light Source.

Environmental groups have been pressuring the Department of Energy to close the reactor since the leak was discovered, even though the total amount of tritium discharged is less than that contained in an emergency exit sign in a cinema.

Opponents of the reactor were able to build on powerful anti-nuclear sentiment on Long Island. Construction of a nuclear power plant there was aborted after residents argued that the densely populated island could not be evacuated in an emergency.

The Department of Energy responded to the leak by firing the consortium of universities that had operated the laboratory. The new contractor had been making some progress in improving its relations with the laboratory's critics in the community (see *Nature* 400, 303; 1999).

But Michael Forbes, the local congressman, who recently switched from Republican to Democrat, has opposed the restarting of the reactor and has inserted language in appropriations bills for three years in a row expressly prohibiting it. Whatever Brinkley and Baldwin told Richardson, the energy secretary must have concluded that the department could no longer afford to spend an annual \$20 million maintaining an ageing facility that local politicians would not allow to reopen.

Colin Macilwain

Science lobby 'ecstatic' after triumph in NIH budget battle

Washington

The US National Institutes of Health (NIH) last week secured a 14.7 per cent increase in its budget for next year, achieving a remarkable triumph after an arduous congressional budget battle. The decision raises the biomedical agency's budget for the fiscal year 2000 to \$17.9 billion.

The House of Representatives and the Senate approved the increase just before adjourning for the year after weeks of struggle with the White House over budget priorities. The sum was part of a \$385 billion, 'omnibus' spending bill that combined five of the 13 bills funding the government.

Congress also last week approved a five-year extension to the research and development tax credit, which creates incentives for private industry to fund research projects.

The \$2.3 billion in new NIH funding, which keeps the agency on track for doubling its budget in five years starting this year, comes with two strings attached. The agency must wait until next 29 September, the end of the fiscal year, for \$3 billion of the overall budget. This is a budgetary device employed by Congress to allow it to appear to avoid spending Social Security revenues.

The NIH plans to handle the problem with 'split funding' — giving investigators part of their award at the time it is granted, and paying the remainder at the end of the fiscal year. Advocates see this as a big improvement on an earlier proposal that would have delayed \$7.5 billion in NIH funding to late September, and forced the postponement of new grants until then.

The \$3 billion deferral is "something that

we can cope with," says David Kaufman, president of the Federation of American Societies for Experimental Biology (FASEB) and a professor of pathology at the University of North Carolina at Chapel Hill.

In addition, the NIH, like other federal agencies, is expected to be subject to a 3.8 per cent across-the-board spending cut insisted on by Republicans as a symbol of fiscal austerity. This would slice \$68 million from NIH's budget, lowering its effective increase to \$2.23 billion, or 14.3 per cent. But the loss is not definite, as the bill allows the president to exempt certain agencies from cuts.

Such provisos were insufficient to dampen the enthusiastic reaction of research advocates. They said last week that the NIH had pulled off a tremendous coup in wresting such a large increase from a budget maelstrom that saw Republicans fighting to find budget savings in all conceivable areas.

"It's an outstanding outcome. We're just ecstatic," says Tim Leshan, director of public policy at the American Society for Cell Biology. Considering that there were those in Congress who felt that the increase "was just too much too fast," the new money is "a tremendous victory," adds Mike Stephens, a FASEB lobbyist.

The bill includes a \$45 million increase in the NIH budget for the construction of extramural facilities, taking it to \$75 million. University administrators and NIH officials had been complaining that the agency's budget was insufficient for what they describe as the decaying infrastructure of US biomedical research facilities (*Nature* 399, 621; 1999.)

Meredith Wadman

Hughes institute unveils top team

Washington

The Howard Hughes Medical Institute (HHMI), the largest biomedical research charity in the United States, has appointed a management team that is expected to shake up the institute's portfolio in the new year.

Gerry Rubin, a geneticist at the University of California at Berkeley, will join the institute on 1 January as vice-president for biomedical research. He will team up with David Clayton, who will become vice-president for science development.

The two appointments were announced last week by Tom Cech, the University of Colorado Nobel laureate who takes over as president of the HHMI in January. Cech said the appointments to replace Max Cowan, the

institute's retiring scientific director, signalled the expanding scope of its activities.

"Max is irreplaceable, so we've decided to divide things up a bit differently," says Cech, adding that the decision reflects the continuing expansion of the institute's portfolio. "If we just planned to maintain what we have, we wouldn't need two people."

The HHMI's annual expenditure has grown rapidly in recent years, from around \$300 million in 1994 to more than \$500 million last year, as the stock-market boom has fed an endowment now worth some \$10 billion. Four-fifths of the expenditure supports an elite of 300 salaried investigators — all biomedical scientists at US universities — with the rest distributed as grants for under-

► graduate and postgraduate education, and for research overseas.

Cech says Rubin will take over supervision of the investigators, while Clayton will be responsible for “strategic planning” of the institute’s activities. Clayton is currently a senior scientific officer at the HHMI and was formerly associate director of the Beckman Center at Stanford University in California.

Rubin will maintain his laboratory at Berkeley, although it will contract sharply once it has completed its current project to sequence the genome of the fruit fly *Drosophila* in collaboration with Celera (see *Nature* **401**, 729; 1999).

“I’ve felt for a while that it is time to take on a more administrative role,” he says. “I enjoy policy, and this is the ideal job because you don’t have to worry about where the money’s coming from.”

Cech is also keeping his laboratory at the University of Colorado. Both Rubin and Cech plan to spend about 20 per cent of their time on research.

Although all three say that no decisions have been made about new directions at the HHMI, Cech suggests that the institute will explore new ways of encouraging top-quality clinical research in the United States, and is likely to build on plans already announced to appoint investigators in bioinformatics.

But they expect that such decisions will be



Clayton (left) and Rubin: expected to shake up management of the largest US medical charity.

made quickly once they take over operation of the institute in January. “I don’t know what will be announced, but I’d hope that the first year of the administration will be an interesting one,” says Clayton.

Cowan, who has steered the institute’s scientific programmes with an iron hand for the past 12 years, says he plans to pursue his keen personal interest in the history of neuroscience during his retirement.

Purnell Choppin, the outgoing president, will remain at the HHMI’s headquarters at Chevy Chase, Maryland, writing an official history of the institute. There are conflicting views as to whether the institute’s secretive and eccentric billionaire founder set it up primarily as a medical research institute or a tax dodge.

Colin Macilwain

Busquin plans white paper to integrate European research

Brussels

The European Union’s (EU) new research commissioner, Philippe Busquin, has promised to publish a white paper (policy document) early next year outlining his plans for the creation of a European Research Area (see *Nature* **401**, 837; 1999).

Speaking to a meeting of the European Parliament’s committee on industry, external trade, research and energy, Busquin said that the paper will address issues such as how to achieve mobility of researchers across the EU, whether researchers’ careers should have a ‘European dimension’, and how to “stimulate the taste for research among the young and promote the participation of women in science”.

Busquin told the committee that he wanted to raise the total research and development spending of EU member states from its current level of 1.8 per cent of gross domestic product towards the 2.8 per cent of the United States and the 3 per cent of Japan.

Keith Nuttall

Genentech pays \$200m over growth hormone ‘theft’

San Diego

The biotechnology company Genentech last week agreed to pay \$200 million to settle a patent infringement lawsuit, over the alleged use of human growth hormone DNA taken 20 years ago from the University of California at San Francisco (UCSF).

The settlement — the largest such payment in a biotechnology case — ends a federal lawsuit that has exposed a seamy side to the relations between universities and industry at the dawn of the biotechnology era.

The university sued Genentech nearly a decade ago, claiming that a midnight ‘theft’ in 1978 of growth hormone DNA was key to the firm’s development of its blockbuster growth-hormone drugs. The company acknowledged receiving the growth hormone DNA, but insisted that it was not used to produce its drugs.

About \$85 million of the settlement will be split equally between three inventors formerly at UCSF — Peter Seeburg of Germany, John Shine of Australia and Howard Goodman of Harvard University — and two collaborators, John Baxter of UCSF and Juan Martial of Belgium.

The settlement also includes a Genentech

contribution of \$50 million towards the construction of a research building on UCSF’s developing Mission Bay Campus. Genentech has the right to name the building, which will cost \$235 million. The remaining \$65 million will go to the university and UCSF.

UCSF Chancellor Michael Bishop says that the settlement “was negotiated in an amicable manner out of mutual respect. The relationship between these two institutions in the past has been collegial and historic. Now, we can continue in the same spirit.” A joint statement by the university and the company noted that the settlement is not an admission of patent infringement.

Arthur Levinson, Genentech’s chairman and chief executive, said that the company “has decided to put this matter behind us and avoid the distraction and uncertainty of another jury trial covering complex patent issues that are based on events that took place nearly 20 years ago”. Genentech will take the \$200 million as a one-time expense during the next quarter, he said.

Last June, in the eyes of many observers, the university nearly won its demand for \$400 million in damages, which it sought to have tripled because of Genentech’s conduct.

After a six-week trial, eight of the nine jurors found that the university’s patent had been infringed (see *Nature* **399**, 512; 1999), but a unanimous verdict was required. This set the stage for a retrial, scheduled for January.

Seeburg, who is now at the Max Planck Institute for Medical Research in Cologne, remains under investigation for possible scientific misconduct 20 years ago. At last spring’s trial, Seeburg testified about how he took the growth hormone DNA from UCSF to Genentech, where he had gone to work after leaving UCSF.

The UCSF DNA was used to produce Genentech’s drug, he testified, and a *Nature* article in 1979 (see *Nature* **281**, 544–548; 1979) contained “technical inaccuracies” which effectively disguised Genentech’s use of the UCSF DNA. This testimony prompted the ongoing probe.

UCSF officials acknowledged last week that the investigating German authorities have requested assistance, adding that the university is in the process of responding. Contacted in Germany, Seeburg said that he was “very relieved” about the settlement, although he denied that he had been required to contribute towards the legal costs incurred.

Rex Dalton & Quirin Schiermeier

China on verge of putting man into space...

Beijing

China has launched its first unmanned spacecraft, Shenzhou, or "God Ship". The flight has made China the third country to launch a craft capable of carrying a person into space, and has been hailed as a milestone towards realizing that dream.

The vehicle was launched with the latest model of the Long March rocket, from the Jiuquan satellite-launching centre in the northwestern province of Gansu, at 6:30 a.m. local time on Saturday (20 November). The dome-shaped craft was in space for 21 hours and orbited the Earth 14 times before touching down in Inner Mongolia.

During the mission, experiments were conducted on remote sensing from space, environmental monitoring, space materials, life science, astronomy and physics. A new land- and sea-based monitoring and control network was put into use for the first time for the launch.

China's secret Manned Spaceflight Programme began in 1992. The first manned

flight was originally planned to coincide with the fiftieth anniversary of the founding of the People's Republic of China in October — no reasons were given for the delay.

The spacecraft and the carrier rocket were developed mainly by the China Space Science and Technology Group. The Chinese Academy of Sciences and the Ministry of the Information Industry were also involved.

The official Xinhua News Agency said the launch would "strengthen the nation's comprehensive national strength, promote the development of science and technology, enhance national prestige and boost the nation's sense of pride and cohesiveness".

Officials say that a number of unmanned test flights are needed before an astronaut can be sent into space, but give no timetable. The Beijing-backed Hong Kong newspaper *Wen Wei Po* said a manned flight was "just around the corner", however.

Major Chinese newspapers did not appear until midday on Sunday, apparently waiting for the return of the spacecraft.

Reports about the launch covered the front pages of almost all major newspapers.

Russia is reported to have helped China with the development programme. The spacecraft weighed about 8.4 tonnes and was capable of accommodating up to four astronauts. It was said to be based on the Russian Soyuz, but with two pairs of solar panels to generate power, and a cylindrical forward module rather than a spherical one.

Moscow and Beijing signed an agreement to train Chinese cosmonauts in Russia after Russian president Boris Yeltsin visited China in 1996.

Two Chinese cosmonauts received training that year at the Star City Space Centre near Moscow in the full programme needed for a manned flight, as well as intensive courses in physics and Russian. The future astronauts are said to have been picked from among China's best fighter pilots.

China launched its first rocket in 1959. The country's first satellite, Dongfanghong, went into orbit in 1970.

Tian Xuewen

... as a new failed satellite launch embarrasses Japan

Tokyo

Japan's space programme experienced a major setback last week when one of its H-II rockets had to be destroyed shortly after lift-off. An engine failure caused the launcher to deviate from its flight path.

The failure led to the resignation of Toshio Okazaki, vice-minister for the Science and Technology Agency (STA), who also took responsibility for the nuclear fuel accident at a uranium-reprocessing plant in Tokaimura in September. But his resignation fuelled criticism of the government for creating a scapegoat for mishaps at STA-related organizations.

The launcher carrying the Multi-functional Transport Satellite (MTSAT) was blown up seven minutes after its launch, on the instructions of the National Space Development Agency (NASDA). Agency officials say that the engine on the first stage of the launcher stopped working about four minutes after the launch, possibly because of a problem with the fuel pipe supplying hydrogen to the combustion chamber.

The failure, resulting in a total loss of ¥34.3 billion (US\$327 million), was a blow to NASDA and its governing body, the STA, which had hoped to improve the space agency's reputation following last year's failure to launch the world's largest telecommunications satellite (see *Nature* 392, 321; 1998).

Three delays to the launch of the H-II

rocket, initially planned for August, added insult to injury (see *Nature* 401, 204; 1999). NASDA admitted after the accident that it did not inspect the main engine thoroughly, because the same type of engine had worked in past missions.

Hirofumi Nakasone, director-general of STA, has promised a full investigation into the failure and a review of space development strategy under NASDA. "This does not mean that [Japan's] whole space development strategy has a problem, but the current system under NASDA clearly needs to be reviewed and made open to the public," said Nakasone.

Politicians from the ruling Liberal Democratic Party called for a drastic reorganization of the country's space programme after last year's launch failure. This view has strengthened after last week's incident, with a greater emphasis on reorganizing NASDA. This would include merging NASDA with the Institute of Space and Astronautical Science (ISAS) under the Ministry of Education, Science, Sports and Culture (Monbusho) after the planned merger of STA and Monbusho in 2001.

But space scientists are opposed to such a move, saying that a merger of NASDA, which is responsible for large-scale, mission-orientated programmes, and ISAS, an institute focusing on basic research, would be disastrous.

"The merger will not make the problems



Ill-fated: the H-II launcher was blown up seven minutes after its successful lift-off (above).

disappear," says Ryojiro Akiba, former director of ISAS and a member of STA's Space Activities Commission. "NASDA may have to shift from its current application-orientated approach to a more technology- and research-based approach, and ISAS could provide support for such a move; but this should not be interpreted as a need to merge the two organizations."

The failure is also a serious blow to Japan's plan to join the satellite business. Rocket System, a commercial satellite-launching company funded by 73 space-related companies, says the failure may affect its contract with two US companies to launch satellites using H-IIA rockets, which are based on H-II launchers.

Asako Saegusa

NASDA

Gender discrimination 'undermines science'

Brussels

Under-representation of women and discriminatory, "old-fashioned" employment practices are undermining efforts to achieve excellence in European science, according to a report prepared for the European Commission in Brussels that was presented to the commission this week.

The report is the first systematic attempt to gather statistics on European women in science and to use them in policy recommendations. Compiled by the European Technology Assessment Network (ETAN), it says that the number of women in senior scientific positions remains low, despite the different cultures and research systems in the member states of the European Union.

In virtually all of the states, for example, women make up less than 15 per cent of full university professors. Data on membership of the world's academies of science reveal what the report describes as an "extraordinary" and undemocratic picture of the institutions that shape science policy (see Table 1).

The report also shows that women are rarely on committees that set science policy. Some private research organizations, such as the governors of Britain's Wellcome Trust and the scientific advisory board of France's Association pour la Recherche sur le Cancer, have no female members. Nor are there any

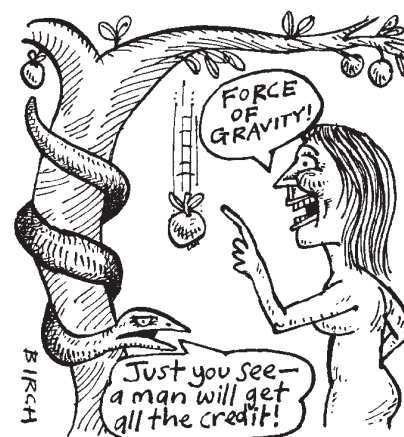
women on the board of governors of the European Commission's own Joint Research Centres.

But one of the report's main complaints is the lack of adequate statistics on gender in European organizations and the difficulty of obtaining data that is comparable across member states.

ETAN is optimistic that the report's official handover to research commissioner Philippe Busquin on Tuesday (23 November), and its consideration by CREST, the commission's main research advisory panel, on Friday will help to maintain the issue's high profile in the commission. The report will then be considered by government officials involved in promoting women in scientific research in member states. Social affairs commissioner Anna Diamantopoulou will also be present on Tuesday.

Chaired by cell biologist Mary Osborn, of the Max Planck Institute for Biophysical Chemistry in Göttingen, ETAN is a group of experts set up last year by the commission to look at Europe-wide issues. Osborn hopes that the report will catalyse discussion at member-state level, and that it will result in a directive requiring employers to keep gender-disaggregated statistics.

ETAN says that member states should introduce legal measures to ensure gender



balance on public bodies and to allow the access to public records needed to examine peer review in the award of grants, fellowships and appointments.

The report says that "old-fashioned" practices characterize some academic institutions. "Reliance on patronage, the 'old boy network' and personal invitations to fill posts cuts across fair and effective employment procedures," it says, adding that scientists' employers tend to be "behind the times in addressing the life/work balance and need to modernize".

ETAN member Agnes Wold, of Göteborg University in Sweden and the author of a study on bias in the award of research grants (see *Nature* 387, 341-343; 1997), describes the report as a "milestone".

"It is very important to show the poor record of [some] countries," she says, adding that women scientists do poorly in countries well known for equal opportunities. "It has nothing to do with childcare; scientists have a special talent — like the military — for keeping out women."

Shaping the commission's sixth Framework research programme (FP6) is another aim of the report. ETAN wants activities and support for networks of women in science, including proposals for 'Eurogroups' and for one-off grants designed to be attractive to women, but not to exclude men.

"We are asking the commission to improve the design of FP6," explains Teresa Rees, the report's rapporteur and a social scientist from the University of Cardiff. The goal, she says, is to make it easier for women to participate. The report says that the organization of science gives men an unfair advantage.

The group, explains Rees, wants "the 'integration of equal opportunities into policies, systems, structures, programmes and ways of thinking and doing'".

"We are proposing a set of strategies towards changing cultures and organizations," adds Rees. "This is a one-off chance to get it right."

Natasha Loder

Table 1 Women members of national scientific academies

Country	%	No. (total)	Society
Turkey	14.6	16 (110)	Turkish Academy of Sciences
Iceland	12.3	19 (155)	Icelandic Society of Sciences*
Norway	11.1	82 (736)	Norwegian Academy of Science and Letters*
Finland	8.0	36 (445)	Finnish Academy of Science and Letters*
New Zealand	7.3	19 (259)	Royal Society of New Zealand
Ireland	6.4	18 (280)	Royal Irish Academy of Science, Polite Literature and Antiquities
Croatia	6.3	9 (142)	Croatian Academy of Sciences and Arts
US	6.2	118 (1904)	US National Academy of Sciences†
Sweden	5.5	19 (347)	Royal Swedish Academy of Sciences
Canada	5.3	48 (899)	Academy III (Science) of the Royal Society of Canada
China	5.1	27 (533)	Chinese Academy of Science‡
Scotland	4.5	54 (1148)	Royal Society of Edinburgh
Austria	4.2	13 (311)	Austrian Academy of Science
Germany	4.0	56 (1378)	All German Academies of Science* § II
TWAS	3.9	20 (512)	Third World Academy of Sciences
France	3.6	5 (139)	Académie des Sciences
UK	3.6	43 (1185)	Royal Society
Denmark	3.5	5 (143)	Royal Danish Academy of Sciences and Letters* ¶
Hungary	3.3	6 (183)	Hungarian Academy of Sciences
India	3.1	21 (679)	Indian Academy of Science‡
Spain	2.7	9 (336)	Royal National Academies of Spain*
Italy	2.6	13 (496)	Accademia Nazionale dei Lincei, Rome*
Ukraine	2.6	5 (192)	National Academy of Sciences of Ukraine
Poland	2.5	5 (199)	Polish Academy of Sciences*
Russia	1.7	10 (600)	Russian (formerly USSR) Academy of Sciences#
Japan	0.8	1 (133)	Japan Academy
Netherlands	0.4	1 (237)	Royal Netherlands Academy of Arts and Sciences* ¶

Ten largest institutions in bold. * includes science and arts; † includes active and emeritus members; ‡ 1994 figures; § in 1991, head count of five regional academies, not including Göttingen and Heidelberg; in 1999, head count of the seven regional academies; II includes foreign and corresponding members; ¶ only the class of sciences has been counted; # 1998 figures.

OECD conference will address safety aspects of biotech

Paris The Paris-based Organization for Economic Cooperation and Development (OECD) is to organize a major international scientific conference in Edinburgh early next year, funded by the UK government, on food safety aspects of biotechnology and genetically modified organisms (GMOs).

The announcement of the conference follows a meeting in Paris last Saturday (20 November) attended by representatives of some 50 consumer organizations and the scientific and business communities. The debate is believed to have marked the first time these organizations have been given a government audience to influence international policies on GMOs.

Pierre Miauton, chairman of OECD's seed certification schemes committee, said the meeting had underlined the importance of consumers in drawing up policy on GMOs.

US societies back federal rules on misconduct

Washington Researchers and university administrators meeting at the US National Academy of Sciences last week strongly

endorsed federal guidelines, released last month, that would standardize the government's approach to allegations of scientific misconduct.

The Federation of American Societies for Experimental Biology and the American Society for Microbiology warmly praised the rules, which define scientific misconduct as "fabrication, falsification or plagiarism" and avoid the use of the term as a catch-all for other forms of professional misconduct (see *Nature* 401, 736; 1999). The societies called for only minor modifications to the guidelines before the final version is implemented by all federal agencies.

Moscow will help build Large Hadron Collider

Geneva The Moscow-based International Science and Technology Center (ISTC) is to provide equipment for the Large Hadron Collider, due to begin operating in 2005 at the European Centre for Particle Physics (CERN) in Geneva. CERN agreed contracts worth SFr12 million (US\$7.1 million) with the ISTC this week, paid for out of two experiments.

The centre is a clearinghouse for the 'conversion' of former weapons scientists, managing financial arrangements and easing the transition between the different economic systems. "Clearly the ISTC has come of age,"

says CERN's research director Roger Cashmore. "That is why the high-energy physics community at CERN has chosen to entrust to the centre major research and development projects of critical importance to the timely construction of the collider's detectors."

Israeli research cooperation with Europe 'paying off'

Jerusalem Israel's participation in the European Union's fifth Framework programme of research (FP5) is already paying off, according to the Israeli daily newspaper *Ha'aretz*. The newspaper quotes Orna Berry, chief scientist at the Ministry of Commerce and Industry, as saying that European commitments to Israeli companies under FP5 stand at 47 million euros. The Israeli government's investment in the programme for 1999 is 40 million euros.

Israel's decision to participate in FP5 was made last year by the then prime minister, Binyamin Netanyahu, despite objections from industry and government. Israel must contribute another 100 million euros over the next three years of the programme. Participation in FP4 was not profitable for Israel. European investments in participating Israeli companies were considerably less than the Israeli government's investment, reports *Ha'aretz*.

Bartfai to lead La Jolla neurological centre

San Diego The Dorris Neurological Research Center at the Scripps Research Institute in La Jolla, California, has named Tamas Bartfai, formerly of Stockholm University and Hoffman-LaRoche, as its first director. Bartfai, 51, will start on 1 January, with further recruitment expected in the near future.

The centre has been funded with a \$10 million gift to conduct studies on diseases such as schizophrenia and Alzheimer's. Formerly chairman of the department of neurochemistry and neurotoxicity at Stockholm University, Bartfai was most recently head of research on the central nervous system at Hoffman-LaRoche in Basel, Switzerland.

Wolff resigns from troubled Arizona observatories

Washington Sidney Wolff, who has since 1987 been director of the National Optical Astronomy Observatories (NOAO) based in Tucson, Arizona, has resigned. Wolff has said that she will stay on until a successor is found.

Her third term as director was due to expire in 2002. The NOAO, which includes facilities at Arizona's Kitt Peak and Cerro Tololo in Chile, has been under pressure to

redefine its mission (see *Nature* 401, 199; 1999). Wolff had been a significant force in plotting the organization's new direction.

Ukrainian biologist on criminal charges

London Ukrainian marine biologist Sergei Piontkovski was formally charged last week with illegal currency operations and "organized crime" by the Ukrainian Security Services in Sevastopol. The charges, which came on a day that Piontkovski had faced a further 12 hours of questioning, relate to the payment of salaries to participants in a research project funded by the European aid programme INTAS and a project funded by the UK government (*Nature* 402, 6 & 110; 1999).

His trial is expected to take place within two months. In an open letter written last week, Piontkovski says that the actions being taken against him were leading him to consider applying for "political shelter in the USA or in Western Europe". His wife has appealed for help with legal fees.

Wellcome targets genetics for biggest cash boost

London Britain's Wellcome Trust has increased its spending on research grants by 70 per cent this year, bringing the total to

more than £400 million (US\$650 million), compared to £235 million in 1998. Research areas receiving the largest cash boost were genetics (up from £17.2 million to £45 million) and neuroscience.

On a percentage basis areas that did best were the public understanding of science — where spending almost tripled to £3.5 million — veterinary science, and tropical medicine and infectious diseases. Trust officials told the *Financial Times* last week that they expect to devote about half of their research expenditure over the next five years to genetics-related research.

Search of Russian ecologist's flat 'illegal'

Moscow The inspection department of the Russian Federal Security Service (FSB) has sent a letter to the Russian Academy of Sciences admitting its officers acted illegally last July when they searched the Vladivostok apartment of a prominent Russian ecologist, Vladimir Soyfer (see *Nature* 400, 300; 1999).

The search warrant said that Soyfer's work — several decades of research on nuclear pollution in the far east of Russia — posed "a threat to the state and military security of the Russian federation". Nikolai Patrushev, FSB director, says he has ordered those who made the search to be punished.

INDIA TREADS WARILY

Licensing difficulties stand in the way of GM crops

p342

BATTLE IN BRAZIL

Farmers and consumer groups take opposing sides in legal fight

p344

FACT FILE

Web and print sources of further information on agri-biotech

p344

ACADEMIES' THINK-TANK

Scientists from the developing world and the West work together

p345

Access issues may determine whether agri-biotech will help the world's poor

Despite doubts in the West about GM crops, many developing countries are keen to exploit them — if they are allowed to do so at an affordable price.

Across vast tracts of marginal land in sub-Saharan Africa, already paltry crop yields are being further decimated by a parasitic weed called *Striga*. The parasite, by a quirk of nature, gravitates towards weak maize, rice and sorghum plants on poorly managed farms.

Transgenic technology, by a quirk of the nature of business, gravitates in just the opposite direction, boosting the already generous yields obtained by the world's wealthiest farmers on the fertile plains of the United States, Canada and Argentina. In the short term, at least, the technology is poised to make these farmers richer and the poorest farmers poorer, as cheap imported grain makes it even more difficult for them to sell their surpluses at local markets.

Noisy arguments between supporters and opponents of genetically modified crops in Europe and the United States have obscured the even deeper ramifications of GM crops for the poor countries of the world, where agriculture is the predominant economic activity, and food security the top political priority.

Disagreements about the environmental safety of GM crops are now spreading to the developing countries. But, for most of them, the more pressing question is how to obtain access to a technology that is being developed, patented and tightly controlled by researchers and corporations in wealthy countries.

Plant scientists believe that the scourge of *Striga* could be lifted if local crop seed was treated with a modern herbicide, such as Monsanto's Roundup. According to Gary Toenniessen, deputy director of agricultural science at the Rockefeller Foundation, such a treatment would involve a minuscule amount of herbicide: as little as 5 grams for every hectare planted. But this treatment could only be done if the crops themselves were genetically modified to tolerate the herbicide.

The New York-based foundation, which has been a powerful supporter of agricultural research in the developing world for decades, believes it will soon find a way to introduce such genetically modified crops into field trials in Africa.



JEREMY HARTLEY/PANOS

To turn a blind eye to 40,000 people starving to death every day is a moral outrage.

But before that happens, Toenniessen notes, plant researchers will have to overcome "all kinds of intellectual property rights issues," concerning not just the gene that confers herbicide resistance, but also the enabling technologies required to get the gene into the crop.

Genetic markers and tissue culture have been used by plant researchers in the developing world for many years with little fuss. It is the arrival of the third and most powerful manifestation of agricultural biotechnology — transgenics, or the transfer of specific genes from one species to another — that is at issue now.

As the first generation of genetic crops enters global agriculture, the stakes for the developing countries are immensely high.

"We are standing at a crossroads," as William Padolina, deputy director of the International Rice Research Institute in the Philippines, told a meeting at the World Bank in Washington. The route taken next, he adds, "will rest on our souls and on our values."

"In the developing world you've got several dimensions to this debate," says Alex McCalla, head of rural development at the World Bank. "One is an extension of the debate in Europe, carried out by some of the non-governmental organizations. The second part is the domestic environmental groups, who see the technology as a threat to biodiversity. The third element is domestic agriculture, which fears it might be left out."

At the World Bank meeting — convened by the Consultative Group on International Agricultural Research (CGIAR), a network of developing-world agricultural research centres, and the US National Academy of Sciences — the degree of polarization in this debate was plain to see.

Klaus Leisinger, head of the Novartis Foundation for Sustainable Development in Switzerland, attacked critics of the technology on moral grounds. "To turn a blind eye to

▶ 40,000 people starving to death every day is a moral outrage," he said, adding that "we have an ethical commitment not to lose time" in implementing transgenic technology.

Leisinger also advised "people who don't have an in-depth view of molecular biology" not to make judgements on issues that require such knowledge. That drew an angry response from Mark Sagoff, a biotechnology critic and ethicist at the University of Maryland, who accused Leisinger of asking that "those who know about biology should talk about biology, those who know about God should talk about God, and then run it all through a cost-benefit analysis".

Sagoff and other biotechnology critics contend that poverty, not low agricultural productivity, is the root cause of hunger in the developing world. Armed with evidence that the world already produces plenty of food, they say that the real problems of poor farmers are lack of capital and infrastructure and the dumping of subsidized produce from Europe and the United States.

It is "misleading" to justify transgenic technology by saying that it will relieve hunger, argues Miguel Altieri, a professor of ecology at the University of California at Berkeley and chair of the committee established by the CGIAR to liaise with non-governmental organizations. "The major technical advances [in transgenics] don't have anything to do with helping the developing world to produce more food."

Few would argue that the first generation of transgenic crops is of much help to poor farmers in the developing world. But a pow-

erful case can be made that these farmers do need better productivity, which transgenics could eventually provide. "The debate about whether the world is running out of food misses the point," muses Derek Byerly, chief economist in the rural development division at the World Bank. The real point is that the economies of the poorest countries depend on farming, for both income and jobs. "To raise incomes you need growth in the agricultural sector," says Byerly. "It is the only way forward."

Byerly thinks that this growth will come from genetically modified crops once other options have been exhausted. "If you are on the yield plateau" — where yield increases by



Staple: the International Rice Research Institute works with transgenic rice in the Philippines.

conventional means have been exhausted — "it is the main way forward". As a result, he says, "developing countries are rallying behind" the technology.

If the potential benefits of GM crops for developing countries are considerable, the risks are harder to pin down (see *Nature* 398, 651–656; 1999). Many poor nations are rich in biodiversity, and feel particularly vulnerable to environmental contamination, however. "Those with the greatest need for this technology have the least resources available to protect the environment," says Ariel Alvarez-Morales of CINESTAV, a research institute in Mexico, where the release of GM maize has been delayed by concerns about its impact on local, wild relatives of the plant.

Many scientists, especially in the United States, detest what they regard as unreasonable attention given to hypothetical risks from GM crops. "This is probably the safest technology that human beings have ever invented," says Nina Federoff, head of the biotechnology institute at Pennsylvania State University and a member of the US National Academy of Sciences. "The countries that need it are the less-developed countries," she says, adding that it would be "tremendously selfish" of developed countries to block it "because of misinformation".

The problem with this assessment, say critics of the technology, is that the developing countries are ill-equipped to handle any health or environmental problems that might arise. Fred Gould, an ecologist at the North Carolina State University, draws an analogy with the 'millennium bug'. The

HELDUR NETOCNY/PANOS

India intends to reap the full commercial benefits

The Indian government took a conscious decision to promote transgenic plant research early this year, but is in no hurry to make a policy on the commercial introduction of GM foods. "We are not rushing with a policy because no engineered crop is commercially grown in the country and no GM food is on supermarket shelves yet," says Manju Sharma, secretary of the Department of Biotechnology (DBT). "Our policy is to continue with research on transgenic crops while simultaneously strengthening our laboratory system to be able to evaluate the safety of GM foods whenever they are ready to be introduced" (see *Nature* 397, 188; 1999).

GM food is hardly a burning public issue in India. Neither the politicians who fought recent elections nor most of the country's farmers, even those who burned Monsanto's *Bt* cotton last year, know much about it (see *Nature* 396, 397; 1998). The issue of GM safety was barely debated even within the

scientific community until last month, when the Indian National Science Academy (INSA) in New Delhi convened a meeting to discuss the relevance of GM foods to India. The meeting was held at the request of the Royal Society of London, which is compiling the views of the academies in the developing world on the subject (see page 345).



Ghosh: 'benefits of GM crops substantial'.

But in the past ten years India has spent US\$15 million on transgenic plant research. Rules and regulations for dealing with genetically modified organisms were framed in 1989, six years before the first field trial took place. Biosafety panels operate at 125 research institutions, and committees are in

place to authorize and monitor field trials.

Most of India's agricultural scientists want the government to embrace GM technology, arguing that it is one of the few options available to provide nutritional and food security for the country's growing population. "India is taking the stand that there is a need for GM plants in world agriculture, not just in India," says a statement from the INSA. The thrust should be on enhancing agricultural productivity using all available tools of advanced science, according to Sharma. "I am keen to introduce GM foods capable of delivering iron and vitamin A — two major nutritional deficiencies in India," she says.

"Transgenics are going to play a major role in Indian agriculture," agrees Prasantha Kumar Ghosh, an adviser to the DBT and secretary of the committee that gives permission for field trials of GM crops. Insect-resistant cotton carrying the *Bt* gene will be a boon to the country, as 45 per cent

problem was quite unforeseen in the early days of computing, he points out, and has been rectified at great expense in the West, but may yet wreak havoc in countries that cannot afford to fix it.

However, governments in developing countries, at least, appear inclined to believe that GM technology can be safely introduced, with China, India, Brazil (at least at the federal level — see page 344), Argentina, Mexico and Kenya among the nations that are strongly supporting it. Public protest is forbidden in China, which is pursuing the technology most aggressively, and is fairly muted in India (see below) and Brazil.

The greatest concern that these and other developing countries have about GM crops is not about safety, but access. “I don’t think that environmental safety or health are really the key questions” for the developing countries, says Calestous Juma of Harvard University, an adviser to the World Bank on the issue. “The key questions are about the structure of agriculture. There is no developing country where this is not an economic debate.”

Researchers as well as farmers in these countries widely perceive the technology to be controlled by a few multinational corporations. Faced with what Ismail Serageldin, chairman of the CGIAR, terms “big laboratories with battalions of lawyers”, they fear that they will never obtain fair access to transgenic technology.

“The last Green Revolution took place mainly in the public sector with open access,” says McCalla. “This one is happening in the



Transgenic maize to help combat the *Striga* weed (above) may win a no-fee licence for use in Africa.

private sector, with all of the intellectual property protection issues that entails.”

“GM technology is being directed by people for whom poverty reduction cannot be a primary objective,” says Michael Lipton, an economist at the University of Sussex, England, who specializes in the economics of poverty. “The key thing is not just to get the technology moving in the developing world,” says Lipton, “but to get it moving for poor farmers in the developing world.”

Some agricultural researchers in the developing countries cannot easily see how that is going to happen. “We don’t have the power to negotiate to protect our rights,” says Behzad Ghareyazie, director of Iran’s Agricultural Biotechnology Institute. “Negotiating power is not in our hands. The innovations are not in our hands. Money is not in our hands.”

The corporations, perhaps shaken by their recent public-relations debacle in Europe, are taking steps to appease these fears. In the case of the *Striga* weed mentioned above, for example, Rockefeller officials say that Monsanto was cooperative at first, although slow to discuss the release of licensing rights. It then transpired that the US corporation had sold its African rights to Rhône-Poulenc. Rockefeller is now confident that it will obtain a no-fee licence for this application of the necessary gene in Africa.

Judy Chambers of Monsanto’s international development programme says that the corporation is “undergoing a paradigm shift: if we want to reach farmers, we have to completely alter our way of thinking”. Monsanto has reached a royalty-free agreement with Kenya, for example, for a transgenic sweet potato, and says it is getting involved in schemes to deliver micro-credit to small farmers.

The corporation finds an unlikely ally in Juma, who observes: “Anyone who depends on knowledge would be foolish not to guard it. Monsanto has 1,500 PhDs at Chesterton [its research campus outside St Louis, Missouri], the biggest concentration in the

Licensing problems are slowing the adoption of GM crops in India.



Hot stuff: transgenic mustard plants in field trials carry genes imported into India.

of all pesticides used in India are for cotton, says Virender Lal Chopra, former chief of the Indian Council of Agricultural Research. “We have technology to create transgenic crops but we don’t have candidate genes. We must go and acquire these,” he says.

Six or seven genes, mostly for insect resistance and herbicide tolerance, have been imported so far. But India has refused clearance for Monsanto to introduce its proprietary genes for crops such as rice and sugar cane, pending an agreement on benefit sharing. In July, the Monsanto Research Centre in Bangalore sought permission to import 33 plasmid DNA constructs from the inventory of its parent company in the United States. Turning down the request, Sharma said it would not be cleared until Monsanto signed an agreement with the government spelling out exactly how commercial benefits would be shared.

GM plants undergoing field trials include cotton, mustard, tobacco, rice, potato, tomato, brinjal, cauliflower, cabbage and chilli, all carrying imported genes. Ghosh says a protein-rich potato, drought-resistant mustard and salt-tolerant rice — crops developed using genes isolated in India — will soon be field-tested. India’s

experience with field trials, he says, has led to an “overall assessment that the benefits of GM crops are substantial, and they may outweigh the risks”.

But not everyone agrees. Vandana Shiva, an activist and president of the Dehradun-based Research Foundation for Science, Technology and Ecology, says that GM crops offer no benefits to farmers. A case filed by the foundation in the supreme court has slowed Monsanto’s schedule for releasing *Bt* cotton in India. Suman Sahai, a geneticist and convener of Gene Campaign, views the technology “as a tool not for eliminating hunger, but to maximize profits of multinationals who control the technology”.

INSA president Govardan Mehta argues that India should not resist GM foods on safety grounds just because some European countries are doing so. “Our main concern is food and nutritional security, and the relevance of GM foods must be seen in this context,” he says.

K. S. Jayaraman

▶ world." The attention the corporation pays to intellectual property rights simply reflects commercial reality, he says. "They aren't worried about Syria," says Juma. "They are worried about DuPont."

And some agricultural researchers report more trouble getting technology out of universities than corporations. "American university researchers are the worst," says the director of one agricultural research centre in Latin America, adding that "they all think they're going to get rich" through licensing agreements.

Even if licences are available at an affordable price, just negotiating through the maze of intellectual property rights on genes and associated technologies is a nightmare for plant researchers in the developing world, according to Richard Jefferson, director of the Centre for the Application of Molecular Biology in International Agriculture (CAMBIA) in Australia. "Many crucial enabling technologies are controlled or limited," he says.



Juma: 'multinationals will reach deals with developing countries.'

Jefferson thinks that most agricultural scientists don't know how intellectual property rights work and that "the scientific community is getting reamed" by businesses that know the ropes. CAMBIA is developing software tools which, he says, will help researchers to work their way through patent databases and find out what claims exist on the genes and technologies they want to use. He laments the way in which the most basic techniques for

genetic manipulation, using agrobacterium transformation or particle bombardment, were developed with public money and then exclusively licensed to private corporations. Jefferson wants publicly funded agricultural researchers to develop methods that will be freely accessible for research and commercial use in poor countries.

Juma sees a different route to such access: he envisages umbrella agreements between

the multinational corporations and developing countries. The corporations will provide access to their technology in exchange for entry to new markets, he says. "Deals will be worked out behind the scenes. Some are probably in the works right now." Following the collapse of a fairly comprehensive agreement with India, however, Monsanto's public position is that it wants one-off agreements for particular genetic solutions.

Agri-biotech information sources

- The biotechnology home page of the Center for International Development at Harvard University, which features extensive discussion of the role of agricultural biotechnology in the developing world, is at: <http://www.cid.harvard.edu/cidbiotech/homepage.htm>

- Full details and proceedings of the meeting, "Ensuring food security, protecting the environment, reducing poverty in developing countries: Can biotechnology help?" — convened by the Consultative Group on International Agricultural Research (CGIAR) and the US National Academy of Sciences in Washington last month — are at: <http://www.cgiar.org/cgnas.htm>

- *World Food Prospects: Critical Issues for the Early Twenty-First Century*, a new report from the

International Food Policy Research Institute, can be found at: <http://www.cgiar.org/iftpri/>

- Rural Advancement Foundation International, a Canadian group concerned about the impact of genetically modified foods on biodiversity, is at: <http://www.rafi.ca>

- The Biotechnology Industry Organization, which represents the biotechnology industry in the United States, is at: <http://www.bio.org>

- A report by the UK Nuffield Council on Bioethics, *Genetically Modified Crops: The Ethical and Social Issues*, can be found at: <http://www.nuffield.org/filelibrary/pdf/gmcrop.pdf>

- Two recent Commentaries in *Nature* have highlighted aspects

of the debate over the use of agri-biotech in developing countries. In "Why Africa needs agricultural biotech", Florence Wambugu, director of the International Service for the Acquisition of Agri-Biotech Applications, argues that there is an urgent need for such technology to counter famine and environmental degradation (*Nature* 400, 15–16; 1999). In "Much food, many problems", Anthony Trewavas, of the Institute of Cell and Molecular Biology at the University of Edinburgh, argues that a new agriculture, combining genetic modification technology with sustainable farming, is our best hope for the future. In the face of "the old enemies of locusts, floods, diseases and pests", writes Trewavas, diversity in technology is "a strength and a necessity, and not a luxury" (see *Nature* 402, 231–232; 1999).

Smugglers aim to circumvent GM court ban in Brazil

A legal impasse over genetically modified crops in Brazil means that GM soya bean seeds are being smuggled into the country from Argentina, especially in the southernmost state of Rio Grande do Sul, whose government has barred their entry.

The resort to contraband by a growing number of farmers reflects the stalemate that characterizes the use of GM crops in Brazil.

The two most commonly produced grains in Brazil are soya and corn (maize). Brazil's annual soya bean production is usually around 25 million tonnes. Corn production is around 34 million tonnes. Rio Grande do Sul, which accounts for less than a tenth of the country's economy, grows about a quarter of its soya beans.

Responsibility for regulating GM



organisms rests with a federal body, the National Technical Commission on Biosafety. Last September, the commission authorized Monsanto to sell genetically modified soya bean seeds that are resistant

to the company's Roundup herbicide. But consumer and environmental groups had the move blocked in the federal court. Monsanto cannot sell the seed until it has been registered at the Ministry of Agriculture, and the judicial question must be decided before that can happen.

Around 45 per cent of the soya fields in Brazil use seeds originally developed by the Embrapa agricultural research institute, the research arm of the Ministry of Agriculture. Embrapa's seeds are adapted to the country's ecological conditions, and the research institute has reached an agreement with Monsanto to jointly develop GM soya bean seeds.

GM crops with added herbicide-tolerance and insect-resistance traits are still being tested in field trials around the

Whatever happens, the interests of the poorer nations will only be protected where they have the technical basis on which to make informed judgements about transgenic crops, and where a substantial, publicly supported research effort is in place to keep abreast of this increasingly private field of research.

"The more information the developing countries have, the better placed they will be to participate in the debate on biotechnology," Juma told the meeting at the World Bank, adding that this sentiment was not universally accepted.

"We can't deny anybody access to this technology, provided they have full information," says McCalla, warning that the developing countries "must not be foreclosed by others" from deciding how to use transgenics.

But for implementation, resources are needed, even as support for publicly funded agricultural research stagnates around the world. The CGIAR network, for example, has enjoyed what one researcher calls "mission creep without funding creep" in recent years, with new missions such as biodiversity conservation making heavy demands on its \$350 million annual budget. Less than 10 per cent of its funds are available for biotechnology research.

"We should all be advocates of as strong a public agricultural research system as we can have," declares Sam Dryden, head of the Colorado-based company Emergent Genetics and chair of the CGIAR private-sector committee. "Sam Dryden and I agree on one thing," says Altieri, the biotechnology sceptic who leads the corresponding committee for non-governmental organizations. "Poor farmers' needs are going to be of no interest to the people who control this technology."

Colin Macilwain

Academies link to map scientific priorities for meeting real needs

Representatives from seven scientific academies — two from the developed world and five from the developing world — are planning to meet early next year to agree on guidelines for the type of biotechnology that is most appropriate to developing countries.

The meeting will be the second — the first was held in London in July — of a working group set up to work towards a common position on the scientific and technological priorities for biotechnology in developing nations (see *Nature* 399, 721; 1999).

The latter are represented on the committee by scientists from the national academies of India, China, Mexico and Brazil — precisely the countries that are engaged in public debate about the relative merits and dangers of the use of biotechnology in crop production — as well as the Third World Academy of Sciences. The developed world is represented by Britain's Royal Society, and the National Academy of Sciences in Washington.

"We are gathering material from the countries concerned to flesh out what is thought to be important for food production in each of them," says Brian Heap, foreign secretary of the Royal Society and an endocrine physiologist who was formerly director of the Babraham Institute, Cambridge.

"We hope to draw together the details in a paper that will, for example, specify the type of trait qualities that will be important for crops in the future. We hope to come to some consensus focusing on what may be called

'phase two' alterations in genetically modified crops, in contrast to the phase one modifications already familiar in developed countries, such as herbicide resistance."

One priority that could be endorsed in such a document, he suggests, is the importance of constructs that increase vitamin A production. "I am also particularly keen to see the problem of anaemia addressed, and in that context increasing the iron content of crops must also be a high priority."

Heap says that the working group is not seeking to diminish the importance of cash crops. "But we are also keen to see emphasis given to such quality traits."

Intellectual property issues have already been high on the working group's agenda. Heap says "we may come forward with recommendations", adding that input from the representatives of the less developed countries will be essential.

"So far we have got quite a measure of agreement and understanding," says Heap. He accepts that in some countries there is strong opposition to the introduction of genetically modified crops, but says that this will be met through the principle of 'subsidiarity' — the idea that individual members will be free to adopt their own local practices.

"Our concern is to make sure that if there is a clear scientific case for the development of these technologies for the less developed countries, and that they will really make a contribution to food security, then we should say that."

David Dickson

Brazilian farmers are fighting back against legal barriers to GM crops.

country, by about 20 companies and research institutes. More than 160 trials have been approved by the federal government, mainly corn and soya bean, but also including such crops as sugar cane, tobacco and cotton. Among other GM plants being researched in Embrapa's laboratories is a staple Brazilian food, beans.

But the Brazilian Institute for the Defence of Consumers (IDEC), which together with Greenpeace took the court action to stop commercial use of GM crops, is adamant that GM products must be labelled as such, and segregated from traditional crops. "We are fighting to have rules that work, like labelling, which is fundamental to guarantee freedom of choice," says Marilena Lazzarini, IDEC's executive coordinator.

And the state government in Rio Grande do Sul has banned GM crops completely. Olivio Dutra, the state governor, is from the left-wing Workers' Party, which is at odds with the federal government in Brasilia on most issues. Besides its traditional suspicion of multinational corporations, the party has justified its opposition to GM soya beans on the economic basis that the country's farmers will be able to export GM-free crops to the European Union.

The state has imported kits for testing GM seeds, and the police have burned any that are detected. Some mayors within the state have declared that GM crops are welcome within their municipalities, but they lack any real power to challenge the state government decision.

Farmers' groups are starting to fight

back to win access to GM seeds. Four agricultural associations, including the powerful Sociedade Ruralista Brasileira, which represents large commodity farmers, and the association of seed producers, have published full-page advertisements in major newspapers defending "the importance of biotechnology to Brazil". The advertisements argue that GM crops have been proven safe in other countries, and that Brazil has introduced a solid legal framework to regulate genetic engineering.

The farmers argue that further delays in the commercialization of GM crops will harm the country *vis-à-vis* other major grain exporters where such crops are already being extensively harvested — the United States, Canada, Australia, Argentina, Mexico and China.

Ricardo Bonalume Neto

Promoting a standard nomenclature for genes and proteins

Sir—The HUGO Gene Nomenclature Committee was pleased to note that *Nature* has, not before time, decided to be more rigorous in requiring authors to use standard nomenclature for genes and proteins (Opinion, *Nature* **401**, 411; 1999). A growing number of journals now realize the contribution they can make to reducing confusion in the published literature and databases.

Our committee provides a searchable list of human gene symbols, including both the approved symbols and many published synonyms or aliases (www.gene.ucl.ac.uk/nomenclature). It is true that nomenclature authorities are finding it hard to keep pace with the overwhelming amount of data. However, the attitude in the research community should not be 'someone should do something', but 'what can I do?'. Scientists, journal editors, database administrators and funding bodies all need to be nomenclature-aware.

It should not be too difficult for most people to find the nomenclature authorities, and the standards and guidelines, for their areas of interest. Far too often, problems arise because editors and authors are ignorant of established standards. Authors have argued against using a previously agreed symbol when describing the same gene in a new publication "because it would detract from the novelty of their results". We have also seen journals publish papers about a "new" gene, whereas if they had contacted us we could have established that the gene had been previously described, but with a different name. Journals are in the best position to enforce nomenclature standards. Researchers are not obliged to submit data to databases, correct existing database entries, or use approved nomenclature—unless they have to do so as a requirement of journal publication.

It was suggested that databases need to do more to help standardize nomenclature. Primary sequence databases (GenBank, EMBL, DDBJ) would have a major problem enforcing standard nomenclature. The volume of data submitted to them precludes all except automated initial checks. It would be possible to check that any gene symbol in the annotation was approved, but ensuring that it was the correct symbol for that sequence would be much more complex. An incorrect, but apparently approved, gene symbol attached to a sequence record is worse than useless.

It is at the level of the secondary, curated databases that much more is being done to ensure that references, sequences and related

data are linked to the approved nomenclature as well as to known aliases. Even with extensive collaboration, this labour-intensive process is under considerable pressure. Recent projects, such as LocusLink at the US National Center for Biotechnology Information, are helping to improve links between related databases, and are bringing the issue of approved nomenclature to the attention of a wider community. The downside of this is more pressure on curators to deal with queries.

Funding bodies should consider whether sufficient funds are being allocated to the curation effort, as without it much effort is wasted trying to find relevant data.

Julia White, Hester Wain, Elspeth Bruford, Sue Povey

*HUGO Gene Nomenclature Committee,
The Galton Laboratory, University College London,
4 Stephenson Way, London NW1 2HE, UK
e-mail: nome@galton.ucl.ac.uk*

Sequencing challenge

Sir—We welcome your editorial on the need to standardize gene nomenclature¹. As you rightly point out, a nomenclature standardized within and between species, and well curated databases, are vital if diverse data are to be integrated.

The Mouse Genome Informatics Nomenclature Committee applauds *Nature's* decision to be more rigorous in requiring authors of papers to use approved nomenclature of genes and proteins. At least four genetics journals require such approval for gene symbols, and others include information on approved nomenclature in their guidelines to authors.

The 1998 "Genetic nomenclature guide"² contains information on the rules for genetic nomenclature in many model organisms, together with websites and contact information. We hope other journals will join the effort to encourage the use of standardized gene nomenclature.

There are numerous articles about the need for approved nomenclature^{3–6}. And two international workshops have been held to bring together researchers on various species and gene families, pharmaceutical interests and database administrators to discuss issues pertaining to nomenclature across species and gene families, and the integration of nomenclature processes between databases^{7,8}.

A major challenge in the coming years will be to address the nomenclature of genes identified by large-scale genomic sequencing. Ideally, genes should be given names that are not only unique, but are also portable between species, and give some indication of relationships within gene families. A third workshop, in May/June 2000 in Breckenridge, Colorado, will

discuss the needs resulting from large-scale genomic sequencing efforts.

Data integration depends on the correct identification of genes and genomic segments. We continue to work with the scientific community to standardize nomenclature within the Mouse Genome Informatics (MGI) databases (<http://www.informatics.jax.org/>) in accordance with the Mouse Nomenclature Rules and Guidelines (<http://www.informatics.jax.org/support/nomen/>). MGI curators work with corresponding authors to determine correct nomenclature before and after publication. This work is done in collaboration with nomenclature groups for human, rat and other species.

Lois J. Maltais*, Ian Jackson†

**Mouse Genome Database Nomenclature
Coordinator, The Jackson Laboratory, 600 Main
Street, Bar Harbor, Maine 04609-1500, USA
†International Committee on Standardized Genetic
Nomenclature for Mice, MRC Human Genetics Unit,
Western General Hospital, Edinburgh EH4 2XU, UK*

1. *Nature* **401**, 411 (1999).
2. Wood, R. (ed.) "Genetic nomenclature guide with information on websites" (Suppl.) *Trends Genet.* (1998).
3. *Nature Genet.* **18**, 89–90 (1998).
4. White, J., Maltais, L. & Nebert, D. *Nature Genet.* **18**, 209 (1998).
5. http://genetics.nature.com/nomen/nomen_article.html
6. Smaglik, P. *The Scientist* **12**, 1 & 6 (1998).
7. Blake, J. A. et al. *Genomics* **45**, 464–468 (1997).
8. White, J. A. et al. *Genomics* (in the press).

What's in a name?

Sir—Your editorial on the state of gene and protein nomenclature struck a chord with those of us recently engaged in compiling the *Oxford Dictionary of Biochemistry and Molecular Biology* (*Nature* **401**, 411; 1999). Formidable complexities face the groups endeavouring to bring order into the situation.

In the first place, one has to distinguish the roles of systematic names and common names. A systematic name aims to provide unambiguous information concerning some aspect of the substance described. In the case of organic chemicals, the property addressed is structure; in the case of enzymes, the enzyme reaction. Which property of a protein would be addressed in the provision of a systematic name—structure, function or origin?

Whichever property is addressed, it is likely that the systematic name would be too unwieldy for everyday use, hence the need for common, or trivial, names. The current fashion for naming proteins following the convention for genes, using three letters followed by a number, proves inadequate in several ways for use as a common name, especially in that it is not descriptive, a property highly desirable in a common name.

Your proposed insistence that authors should as far as possible indicate all related

names for the protein they are describing is indeed a step in the right direction.

Tony Smith

Oxford Dictionary of Biochemistry and Molecular Biology, 46 Stanley Hill Avenue, Amersham HP7 9BB, UK

Cooperative efforts around Lake Tahoe

Sir—The proposed Round Hill facility of the Desert Research Institute in cooperation with the University of Nevada, Reno, will be complementary to the University of California at Davis facility on the California side of Lake Tahoe, and will certainly not be a “duplicate” as suggested in the News story “University confronts new rival across Lake Tahoe” (*Nature* **400**, 806; 1999).

Our Nevada facility will focus on atmospheric, ecological and hydro-geological research on processes in the forest environments surrounding the lake, whereas most Davis work has dealt primarily with processes in the lake. The Round Hill facility will also provide a location for integration of science, management and public policy.

The planning for Round Hill has been open, with details discussed in a series of public meetings since 1997. A memorandum of understanding was signed on 11 August 1999, in which the Desert Research Institute (DRI), University of Nevada, Reno (UNR), and Davis agreed to work with federal and state agencies to develop and implement a research agenda. This is not “competition” by Nevada scientists. The atmosphere is much more cooperative than suggested by the News story.

The notion that UNR and DRI scientists are newcomers to Tahoe basin research is not true. Nevada researchers have been studying the basin’s ecosystems for decades, often in collaboration with Davis scientists. The Davis Tahoe Research Group, under the leadership of Charles Goldman, has also made significant contributions to our understanding of processes within the lake.

The proposed research facilities will be of benefit on two counts. First, the ecosystem is so complex that two research centres, one in Nevada and one in California, should speed the process of tackling the lake’s environmental problems. Second, the US system of research funding is based on competition so that the best research is conducted. Although we anticipate collaboration on many projects, additional scientists competing for funds will increase the quality of research in the end.

Our goal is to provide excellent science that will allow public policy-makers to define and develop effective environmental management strategies. Nevada scientists

support the excellent collaborative research programme developing at Lake Tahoe, which includes faculty members from Davis as well as researchers from national and international institutions. Our combined efforts will lead to an environmental management strategy to sustain the lake’s beauty for the foreseeable future.

Stephen G. Wells*, James S. Coleman*, Joseph N. Crowley†, Kenneth W. Hunter Jr†

**Desert Research Institute, 2215 Raggio Parkway, Reno, Nevada 89512-1095, USA*

†Office of the President, University of Nevada, Reno, Nevada 89557, USA

Rex Dalton, the author of the News article, stands by his story as an accurate description of the issues and competitive situation surrounding Lake Tahoe research. — Editor, *Nature*

Precautionary approach to risk assessment

Sir—The meeting of the World Trade Organization (WTO), which opens on 29 November in Seattle, Washington, will be a focal point for discussions about the social, economic and environmental implications of trade. Science has been given a central role in mediating disputes about the safety and environmental impact of new technologies. Our research indicates that the WTO will need to be both more rigorous and more precautionary in its use of regulatory science if it is adequately to address issues of sustainability.

First, there is the issue of product safety. Conventional probabilistic risk assessments play an important role in WTO rulings about the safety of new products, such as genetically modified (GM) foods and hormones used in animal production [see Briefing in this issue, pages 341–345]. In contrast, ‘precautionary approaches’ are often thought of as deviating from sound science.

Yet current risk assessments can only characterize some of the potential outcomes of the use of new products. Precautionary approaches, by acknowledging our incomplete knowledge of possible outcomes, and addressing the huge variation in the results of risk assessments, may actually enhance the rigour of scientific assessment. We developed this argument in *The Politics of GM Food: Risk, Science and Public Trust* (see www.gecko.ac.uk).

Second, there is the question of burdens of proof. WTO regulations and other international trade rules increasingly assume that new products are safe until proven otherwise: the burden of proof falls heavily on those who are worried about unforeseen or untested safety and environmental issues.

The debate about the safety of the

bovine somatotropin (BST) growth hormone is a recent example. It has fallen to the European Union (EU) to provide evidence to support its doubts, rather than the producer of BST to demonstrate its safety. So the burden of proof falls on the regulator rather than the proponent of the technology. We believe that the burden of proof should be re-balanced through enhanced and transparent testing of new products, similar to the current method for testing and approving drugs. Likewise, there is a need for better scientific monitoring of the effects of new technologies once in use. Hormone-disrupting chemicals exemplify the issue of inadequate monitoring.

Assuming that products are safe until proven otherwise may lead to what can be described as ‘soft disasters’ — large-scale health and environmental problems that emerge slowly but at high cost to society. Such disasters mostly occur because excessive faith has been placed upon limited data about the safety of a product or process, ignoring many possible eventualities where there is little or no information.

It is now generally accepted that the assessment of risk in different social contexts can produce different — but equally valid — results based in science. Alternative assumptions, for example, are often adopted in different countries, partly as a result of varying social, economic and institutional conditions. But this calls into question the WTO’s apparent assumption that the application of ‘sound science’ will lead to a single ‘scientific’ answer to complex questions of risk and safety. Further, to expect a uniform pattern of associated regulatory decisions is inconsistent with the well-established insights from risk assessment.

Such a desire for single, definitive answers is likely to generate increasing tension in the WTO, and undermine public confidence in its decisions. Public confidence will be vital if the advantages of liberalized global trade are to be sustained. A more effective way forward would be to pursue these issues through a precautionary approach within multilateral environmental agreements. Liability regimes also need to be strengthened as a safety net for those affected by ‘soft disasters’. In this respect, the EU’s recent extension of its strict liability laws to include agricultural products is to be welcomed.

These comments are based on research carried out within the Global Environmental Change Programme of the UK Economic and Social Research Council (ESRC).

Alister Scott*, Andy Stirling, Nick Mabey, Frans Berkhout, Chris Williams, Chris Rose, Michael Jacobs, Robin Grove-White, Ian Scoones, Melissa Leach

**Corresponding author: ESRC Global Environmental Change Programme, Mantell Building, University of Sussex, Brighton BN1 9RF, UK*

Mastering the art of the skies

The observations of Virgil and of a Quaker chemist converge on canvas.

John Constable's Skies: A Fusion of Art and Science

by John E. Thornes
University of Birmingham Press: 1999.
288 pp. £40

Euan Nisbet

Clouds have an ethereal beauty, obvious yet evanescent, that transcends the imagination. Without help, art cannot capture this loveliness; even the Renaissance giants failed. Twice have natural science and art been interwoven to sense the air — once in classical times and again in Hanoverian England. On each occasion, the fusion of careful observation with artistic insight has created beauty out of natural loveliness.

First came Aratus, a Greek scientist-poet who inspired large parts of Virgil's loveliest work, the *Georgics*. Aratus' *Phaenomena* was also cited by Saint Paul in Athens (*Acts* 17.28). The work is based on observation, and remains valid — even though the Earth has moved on (nowadays there is a pole star that was not there for the Greeks). In *Phaenomena*, Aratus described "clouds that look very like fleeces". For Virgil this becomes, "[The Moon] no more glimmers obnoxious to her brother's rays; nor fleecy clouds float lightly through the sky".

Science conspired with art to capture clouds a second time through Luke Howard, a saintly Quaker. In the 1790s, Howard came to London from Manchester. Like his close friend John Dalton, he was a chemist, and he became wealthy by helping to found the modern pharmaceutical industry. In a paper to the Askesian Society, Howard adapted Virgil's language to classify clouds: nimbus, stratus and cumulus, while Aratus' 'sheep's hair' becomes 'cirrus' (a curl of hair).

Howard was supported by another close friend, Thomas Forster, author of *Researches into Atmospheric Phaenomena* and *Notes on Aratus*. In the second edition of Howard's own work, *The Climate of London* (1820), the title page of volume one has a Latin pun on Aratus, while volume three has a twist on Virgil's fleece. From his travels to Europe, Howard brought back Horace de Saussure's 'cyanometer'. This, an index device for measuring the colour of the sky, is vital technology for the painter.

Just as Aratus had inspired Virgil, so in Regency England Howard triggered genius. In Percy Shelley's "The Cloud", based on Howard's cloud classification, "... the Moon glides glimmering o'er my fleece-like floor"; John Keats speaks of a "mid-day fleece of clouds" in *The Fall of Hyperion*. In distant Weimar in Germany, Goethe saw the cirrus



Curles of hair: Constable's "Study of Cirrus Clouds", the result of intense research into the phenomenon.

fleece, too. He commented: "But Howard gives us with clearer mind, the gain of lessons new to all mankind ... As clouds ascend^{stratus}, are folded^{cumulus}, scatter^{cirrus}, fall^{nimbus}", Let the world think of thee, that taught it all" (from "In Honour of Howard", translated by George Soane and Sir John Bowring, in *Luke Howard 1772–1864* by D. Scott; William Sessions, 1976). Not only poets relished the new meteorology. In *Emma*, Jane Austen's observation of the weather is acute: the beauty of the work grows from the delight in the science.

John Constable put Virgil and Howard into paint. A country windmill's son fascinated by the sky, Constable bought a copy of Forster's *Atmospheric Phaenomena*, and in it studied Howard's cloud classification. Before Constable, no artist, not even the great masters, had fully come to terms with the problem of the sky. The Dutch masters and Thomas Gainsborough, whom he studied carefully, came closest, but Constable succeeded. He rented a house in Hampstead village, just north of London, and set off onto Hampstead Heath with his eyes set upwards. Perhaps, as he stared at the sky, he was quizzed by a fellow villager, Keats. A comment Constable made about his patron, Archdeacon John Fisher — "a man to whom nature does not spread her volume or utter her voice in vain" — aptly describes Constable's own spirit. Through Fisher's London home in Keppel Street, later occupied by Constable himself, passed masterpiece after masterpiece.

John Thornes splendidly shows us the workings of art's master of the sky. *Constable's Skies* is a gorgeous work, filled with

colour, pleasing to the eye and mind. This is a book of beauty and rigour, part essay on art, part text on the science of the weather, part catalogue. Thorne treats us to much wisdom, and carefully lays out the fundamental physical basis of the phenomena painted by Constable. There is a fascinating essay on the rainbow, and much detail on the colour of the sky.

The close links between the art and poetry of the time are well drawn. The core of the



Howard's "Cloud Sketches", from the paper presented to the Askesian Society in 1803.

book is the study of Constable's skill as it evolved, guided by scientific insight and his own discovery. Thornes chronicles the growth of the painter's power, mapping out both his original genius and his dependence on Howard and Forster. By contrasting the dates of Constable's art with Howard's daily weather record, Thornes demonstrates the basis of the art: accurate, true observation. Only rarely did Constable deviate, most strikingly in a late masterpiece of Salisbury Cathedral under a violent sky with a physically improbable rainbow. But in this painting, allegory ruled the day: what Constable, referring to Noah, called the "Mild arch of promise", is guarding his mild and threatened Church.

Keats and Wordsworth agreed that Isaac Newton had destroyed all the poetry of the rainbow by reducing it to a prism. Indeed, Caspar David Friedrich, urged by Goethe to improve his cloud painting by studying Howard, refused for this reason. But if Newton broke the waters of the rainbow, it was to allow Howard and Constable to bring to birth a brighter charm.

Keats, ambivalent, saw this: the study of clouds and seasons delighted him. On Sunday, 19 September 1819, not long before his own mortal illness, he walked through Winchester, past the house where the creator of *Emma* had died. Out in the countryside, he saw barred clouds over a warm autumn land. On his return, we can imagine him, building the "Ode to Autumn" in his mind, in the cathedral near Jane Austen's fresh grave, listening to the wailful choir at evensong. Although dark winter comes, science, too, has its seasons of mists and mellow fruitfulness.

To name something is to begin to understand it. Gainsborough's skies have charm, while Constable and his little-remembered equal Copley Fielding bring beauty out of understanding. Joseph Turner enters another world. But in our age, science is now more likely to be the enemy than the friend of art, or to be ignored. The fleece is a dead sheep in formaldehyde, not a sky of lovely cirrus tufts.

Soon after this book arrived, as I walked home under a setting autumnal sky, clearing after a westerly gale but still dark to the east, the bright equinoctial sun lit a perfect, full, double rainbow. Shelley's house is barely a mile away, but students, deep in their inner landscapes, were hurrying home to beer, TV or study, heads down. Only a child on a bike, excited, urged on his laden mother — the pot of gold was right over their house. With his friends Aratus and Virgil, Forster and Howard lighting the chiaroscuro, Constable found that pot, although he moved the bow. ■

Euan Nisbet is in the Department of Geology, Royal Holloway College, University of London, Egham, Surrey TW20 0EX, UK.

More views on the large scale

Fire in the Sky: Comets and Meteors, the Decisive Centuries, in British Art and Science (new in paperback)

by Roberta J. M. Olson & Jay M. Pasachoff
Cambridge University Press, £18.95, \$29.95

The Panorama

by Bernard Comment
Reaktion Books, £45

Personal views of an impersonal world

Meaning in Technology

by Arnold Pacey
MIT Press: 1999. 264 pp. \$27.50, £16.95

Howard P. Segal

Meaning in Technology is Arnold Pacey's latest effort to explore technology in both historical and contemporary contexts. Where *The Maze of Ingenuity* (1975; revised 1992), *The Culture of Technology* (1983) and *Technology in World Civilization* (1990; all MIT Press) dealt with the relationships between technology and such large forces as ideas, culture, politics, economics and geography, this new work considers the relationship between technology and the individual. Yet the author's rejection of technology as being 'value-free', his concern with more humane forms of modern technology and the inclusion of his own experiences and reflections connect *Meaning in Technology* to its predecessors. So, too, does the conversational style that Pacey, an associate lecturer at the Open University in Milton Keynes, brings from the classroom.

In his new book, Pacey's insistence that personal experiences provide 'private meanings' in — and about — technology that are no less valid than the official 'public meanings' creates both the strengths and the weaknesses of the book. His descriptions of many private meanings, while a worthwhile exercise, will not be a revelation for most students of technology. To be sure, Pacey never quite claims complete originality, and generously (sometimes too generously) quotes from others to make his points. Yet he implies throughout that this is pioneering research.

Take, for example, Pacey's visit in 1951 to London's centennial celebration of the Crystal Palace Exhibition. He laments that, in treating technology as applied science, the 1951 exhibition made technology secondary to pure science rather than its intellectual equal. Having at the time uncritically accepted this 'public meaning' of technology, Pacey has subsequently come to reject the equation of technology with applied science

and now wants his readers to do the same. Yet this equation has long since been rejected by virtually all students of technology. In fact, the very field of the history of technology came into being in the late 1950s and early 1960s just because many historians of science, with similar biases, deemed the history of technology an unexciting area of scholarship, best kept apart from the more elevated history of science. There is nothing wrong with making the same point in 1999, but there is nothing pioneering about it either.

Later parts of the book effectively show how engineers, mathematicians, craft workers and consumers experience technology in their work and leisure. Pacey demonstrates how individuals' sense of purpose and meaning in their lives can affect the shapes and applications of the technologies they use. He rightly contends that technology in itself does not generate a hierarchy of meanings and that such values as "empathy, cherishing and charity" are consequently as important as objectivity and detachment. Pacey therefore advocates a more 'participatory' and less 'detached' style of technological activity.

The book is most interesting in its examination of music as a source of technology. Increasingly, Pacey claims, we utilize machines and other technologies in much the same way as we have long utilized music and musical instruments, "to interpret the world and give it meaning". The sounds produced by machines generate meaning, the rhythm of machines is music to many workers, and the experience of music for participants and listeners alike is an antidote to the overly mechanistic world we are ordinarily expected to embrace.

Aware that his highly personal perspective on technology and its meanings could lead to charges of sentimentality and soft-mindedness, Pacey defends himself by arguing that "we should acknowledge the existence of feelings, not that we should indulge them". He rightly insists that his and any other explorations of feelings be conducted with the utmost rigour, logic and rationality. Yet it is exactly such rigour that is missing from the book.

Finally, Pacey just as firmly rejects 'causal models' of technological development that leave out human purpose, expectations and relationships. As with so much else in the book, the 'technological determinism' that Pacey condemns has been condemned by technology students for decades. Nevertheless, *Meaning in Technology* is a sensitive, sensible critique of outdated and ahistorical notions about technology as a suprahuman force transcending individual experiences and feelings. ■

Howard P. Segal is in the Department of History, University of Maine, Orono, Maine 04469-5774, USA.

A friend of one's own

Imaginary Companions and the Children Who Create them

by Marjorie Taylor
Oxford University Press: 1999. 215 pp.
£18.95, \$25

Paul L. Harris

Books on cosmology and biology are popular among non-specialists, but psychology — especially child psychology — has rarely translated laboratory science into material that would interest the general reader. The lay reader of child psychology is typically a parent looking for advice or reassurance, not a disinterested adult seeking food for thought. Marjorie Taylor's book about children's imaginary companions is an unusual attempt to reach both audiences. For the parent seeking reassurance, she provides a judicious review of a wide range of findings showing that children who create and sustain an imaginary companion are not suffering from any obvious clinical disorder. At the same time, for the disinterested reader, she describes many vivid specimens of the phenomenon and engages in enough conceptual analysis to show that this intriguing aspect of children's early fantasy raises

fundamental questions about human imagination.

In defining imaginary companions, she excludes 'transitional' objects — the stuffed toys and blankets that children trail after them — on the grounds that they are not attributed any human qualities. She concentrates instead on those acts of pretence where, for a sustained period, children either conjure up or impersonate a human being or animal. This latter type of companion is not without its domestic complications. The child may remain on all fours, reply to all requests with "woof, woof", lap up food from a dish and even urinate by raising one leg. More typical, and a bit less trying, is the child who invents a somewhat demanding invisible friend — who has to be assigned a place at the dining table or have the television left on if the family goes out.

By interviewing three- and four-year-olds and their parents, and focusing primarily on those cases where the two sources agree, Taylor concludes that the phenomenon is widespread: more than one child in four invents an invisible being and one in ten engages in impersonation.

Why do children create such companions? Taylor's answer is not unsympathetic to over a century's worth of clinical interpretations. Almost all of these assume that it is emotion that drives the child's imagination. The child who creates an imaginary companion is said to be staving off loneliness, seeking out a powerful ally, nurturing a weaker dependent, displacing criticism away from the self or mastering various fears.

Yet there are several problems with this line of thought. It is post hoc — it is always possible to think of some emotional need that a relatively human-like companion

might meet. It cannot explain the finding that children who have lived through an emotionally charged experience, such as abuse or abduction, are less inclined to engage in pretend play than their normal peers. Finally, it does not account for the sheer perversity of some children's companions. Take the three-year-old who had an imaginary pony: when she was taken to

an actual horse show, the child was upset that the pony was 'not there'. Apparently, the child could not — or would not — simply pretend that the pony was there. It requires a heavy dose of theoretical credulity to suppose that some hidden need for separation or absence was at the root of this self-denying ordinance.

Nonetheless, Taylor is surely right to emphasize that imaginary companions engage the child's emotions. What emerges forcibly from her intriguing descriptions is that, even if the emotions do not drive the imagination, the imagination definitely drives the emotions. Indeed, one of the most interesting issues raised by children's imaginary companions is why their emotions are so readily activated by what is, after all, a mere fantasy — why is the three-year-old moved by the pretend absence of a pretend pony?

Taylor rightly insists that the answer is almost certainly not that children are confused about the distinction between fantasy and reality. A large body of findings shows that they grasp this early on. In any case, a similar activation of emotion by mere fantasy can be observed every night in our theatres and cinemas. Presumably, adult audiences who are moved by such fictions are not confused about their status. The phenomenon of imaginary companions raises important, but neglected, questions about adult as well as child psychology. Taylor successfully combines a balanced review of a century of research on the phenomenon with a sensitivity to some of those wider issues.

Paul L. Harris is in the Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK.

Sniffing the world of olfaction

Electronic Noses: Principles and Applications

by Julian W. Gardner and Philip N. Bartlett
Oxford University Press: 1999. 264 pp.
£55, \$100

Vincent Dusastre

Patrick Suskind's *Perfume* (Penguin, 1989) is a poignant and disturbing novel about the crucial importance of our sense of smell to everyday life. Although our olfactory system is generally not as acute as that of Grenouille, the 'nose' in Suskind's book, it deeply influences our social interactions, behaviour, emotions and perception. People like Grenouille are pretty rare, but these so-called professional human noses have been used for centuries to assess the quality of food, drink and perfumes. With the recent advent of 'electronic noses', supposedly capable of mimicking our olfactory system, their jobs might well be on the line.

BUBBLES/ANGELA HAMPTON



Would you like sugar? More than one child in four invents an invisible being.

Historically, the idea of how odours are detected was first suggested in the 1920s, but the concept of creating an electronic nose as a clever device consisting of a chemical-array sensor system for odour classification did not emerge until the early 1980s. Since then, this area of research has undergone rapid and tremendous development.

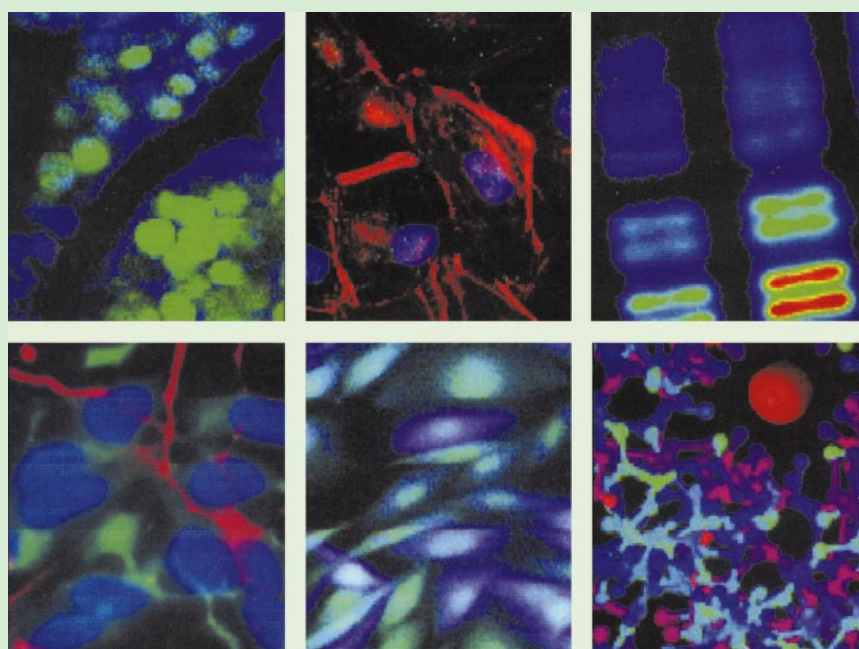
Electronic Noses is a fascinating and instructive book on the way our noses work, the progress made in the artificial detection and recognition of complex odours, and how this technology will be used in the near future. The authors' answer to the question, 'Why bother to try to recognize odours artificially?' is quite straightforward. First, from an academic point of view it is extremely challenging to try to figure out the biological and chemical mechanisms underlying our poorly understood sense of smell. Second, and more importantly, these olfactory machines are useful for routine industrial applications because, in contrast to humans, they are objective, reliable and do not get tired.

Although this is a technical account written mainly for an interdisciplinary scientific readership, the current and potential applications should appeal to anyone with interests from wine-tasting and food quality to environmental issues. This is a much-needed, clear and comprehensive review of the field of artificial olfaction as well as a basic treatment of the wide-ranging areas underlying its science and technology. As Julian Gardner and Philip Bartlett point out, our understanding of human odour detection and how we go about trying to mimic it coincides with a crucial time when instruments for doing this are starting to affect our everyday life.

After its introductory background, the book deals with the many requirements needed to make an actual device. The complex nature of the problem and the need for advanced technologies explains why the commercialization of these noses has only just begun. It is now possible to design devices capable of performing individual tasks 'similar' to our own olfactory system. This is a result of recent developments in microsensor technology, which have led to the creation of inexpensive integrated chemical sensors, together with applications specific to microprocessing devices and a better understanding of artificial intelligence.

Electronic Noses is not exactly an entertaining read. At face value, it is a collection and summary of research studies in the field, many of them from the authors' own work. However, it is much more than this in being, I believe, the first comprehensive and up-to-date overview of the field.

Although the authors — a chemist and an engineer — appear to be on firmer ground when dealing with their own fields of expertise, they clearly and convincingly describe the functioning of our biological olfactory



Cellular variations

Musically minded scientists Ira Melman, Richard Flavell and friends have turned to what they call biorock. The lyrics below are from one of the songs on their rock band's CD, *Cellmates Volume 1 Number 1*. (\$15, or £12 plus shipping. Apply to richard.flavell@yale.edu or ira.mellman@yale.edu)

"Kits-R-Us" (R. Flavell)

Way back in history they used to make up their ATP.

Pour out buckets of SSC, but screw that, suckers, that's not for me!

Now all that stuff just takes too much time, And it's my career that's on the line!

They cost some dough, let my PI whine Cuz I'm headed for the stars this time!

You buy it in one carton, it's got everything you need.

A few instructions to start on, can't fail to succeed.

Dial 1-800-Kits-R-Us (free pipettes with every order)

Dial 1-800-Kits-R-Us (free pipettes with every order).

With just three tubes, red, white and blue, There's not too much for me to do.

Just mix it up, pipette that too, no need for fancy witch's brew.

And if perhaps the buffer's gone, just toss the kit, you can't go wrong!

You save so much time, science speeds along.

That's why I have time to write this song!

processes for the non-specialist. I was disappointed that the historical and philosophical introduction to our sense of smell was not developed beyond the first two pages of the book, but, from the editors' perspective, it would perhaps have been beyond the scope of the book and the expertise of the authors.

But, if you are a gourmet and enjoy eating and drinking as much as I do, the chapter on the mechanisms of human perception of odour and flavour combined with the presentation of our biological olfactory machinery and its links with the brain will not only stimulate your intellect and olfactory neurons, but will also allow you to describe odours more precisely. You should, nevertheless, be warned that because of its technical aspects, *Electronic Noses* might well be difficult to digest in one go. Yet the final three case studies described, on the identification of perfumes, the freshness of fish and taints in beer, demonstrate the authors' desire to make

the book accessible to a wide readership.

The evolution of this technology depends on advances in many disciplines, and, as such, the future of the field is hard to predict. However, the potential market is so enormous and diverse that we can expect intense interest at least in the near future. Although our sense of smell has long been regarded as the least refined of our senses, and however confident the authors are about the rapid evolution of this technology, there is still a long way to go before such devices can accurately reproduce our subtle perception of flavours.

Many articles about this technology have appeared in the specialist literature, but it has not so far attracted much public attention. Hopefully, this comprehensive and thorough discussion of artificial olfaction and its fascinating interdisciplinary character will do just that.

Vincent Dusastre is an assistant editor at Nature.

Two millennia of animal spirits

An ancient theory about nerves proved surprisingly hard to dislodge.

Ian Glynn

I think it was E. D. Adrian who said that in many lines of research there are two crucial steps. The first is to find the right question; the second is to find the right answer to it. Sometimes the interval between the two is quite short. Having asked why the apple fell, Newton didn't take long to produce the theory of gravitation. But the interval between asking and answering what passes along nerves that enables them to produce their effects was more than two millennia.

It was Alexandrian physicians of the third century BC who first asked the question about nerves, and they also attempted an answer. It was, they said, the flow of animal spirits along motor and sensory nerves that led to movement and sensation. As animal spirits (*pneuma psychikon* in Greek; *spiritus animalis* in Latin) were thought of as weightless, intangible and invisible, to say that nerves produced their effects by a flow of such spirits was to say little more than that something undetectable was passing along the nerves — which was of course true. But that doesn't explain why a theory that was neither testable nor helpful survived until the seventeenth century AD. Its survival was the result of its espousal, in the second century AD, by Galen — not so much because of his impressive contributions to physiology and medicine, as because his belief in a single, all-wise, all-just God or Creator made his medical writings almost canonical throughout the Christian and Muslim worlds.

By the seventeenth century, ad hoc, incorporeal spirits had reached their 'best before' date. Descartes regarded animal spirits as a real fluid, and compared the actions of nerves — which he believed contained valves — to the action of the hydraulic pipes controlling machinery in the royal gardens. He said it was the distension of muscles by an inflow of animal spirits that caused them to shorten. But nerves don't contain valves, and Jan Swammerdam, in Amsterdam, showed that muscles don't increase in volume when they shorten. Giovanni Borelli, professor of mathematics at Pisa, rejected the idea that any substance passes along nerves and suggested that what is transmitted is "a commotion". Nerves, he said, were "canals filled with spongy material like elder-pith ... moistened with the spirituous juice of the brain ... and saturated to turgescence", so they could transmit a "concussion or undulation". The basic notion of a transmitted commotion was right; the colourful analogy of concussion through a column of fluid was not.

Descartes regarded animal spirits as a real fluid, and compared nerves to hydraulic pipes.

Discovery of the nature of the commotion had to await an understanding of current electricity — the work of Luigi Galvani, Alessandro Volta and Michael Faraday.

In the middle of the nineteenth century, in Berlin, Emil Du Bois Reymond showed that the voltage across the surface membrane of an inert nerve fibre — the resting potential — was transiently diminished when the fibre was stimulated. What is more, the fall in the transmembrane voltage — the action potential — occurred first at the point of stimulation and spread in both directions along the fibre. Could the action potential be the message that the nerve transmitted? If it were, the message and the action potential would necessarily travel at the same speed. In 1850, Hermann Helmholtz stimulated the nerve of a frog muscle at different distances from the muscle and compared the delays before the muscle contracted. He calculated that the speed of the message that initiated contraction was about 27 metres per second. Eighteen years later, Julius Bernstein, a former pupil of both Du Bois Reymond and Helmholtz, overcame difficulties caused by the great inertia of the galvanometers of the day and showed that the speed of the action potential was not significantly different.

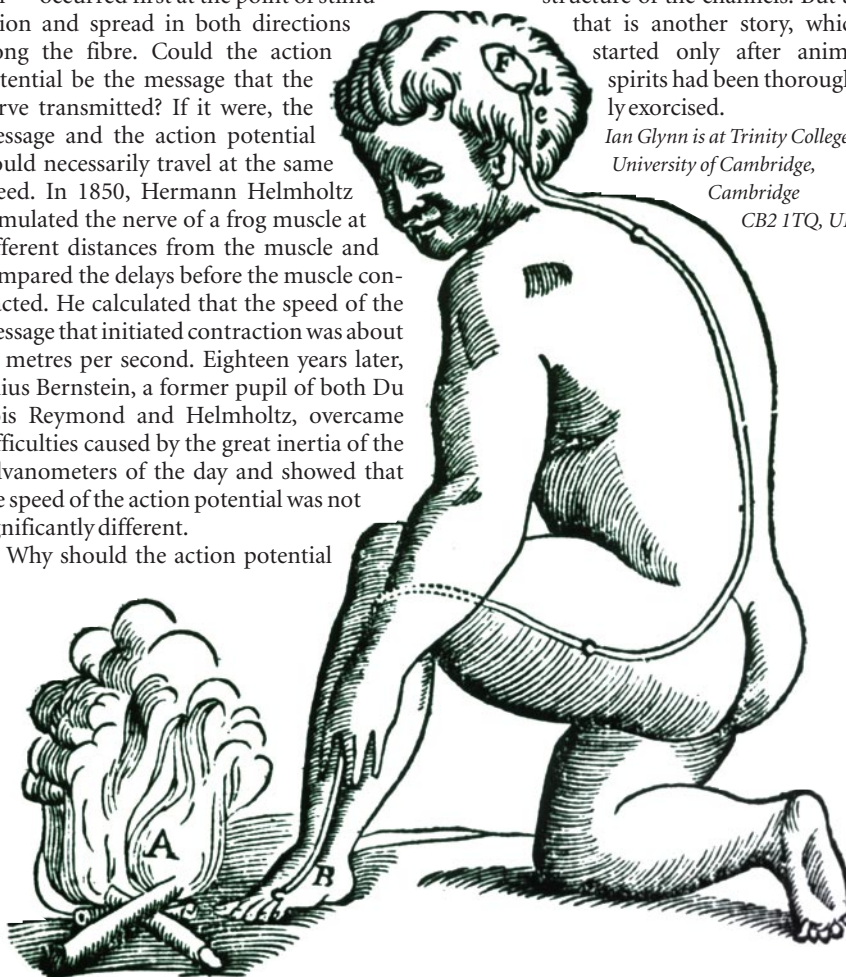
Why should the action potential

move along the nerve? In the last year of the nineteenth century, Ludimar Hermann, in Königsberg, pointed out that, at the junction of active and resting regions of the nerve, local electric currents would tend to reduce the voltage across the membrane adjacent to the active region. If this reduction excited that adjacent region, local currents would again be generated, exciting the next bit, and so on. In other words, the action potential would be propagated along the nerve.

The twentieth century, and particularly the work of Alan Hodgkin and of Andrew Huxley, would prove Hermann's hypothesis, would show that the action potential involves not just a diminution but also a change in sign of the voltage across the nerve membrane, would explain how the subtle behaviour of voltage-sensitive channels selective for sodium or potassium ions gives rise to the resting and action potentials, and would begin to relate that subtle behaviour to the molecular structure of the channels. But all

that is another story, which started only after animal spirits had been thoroughly exorcised.

Ian Glynn is at Trinity College, University of Cambridge, Cambridge CB2 1TQ, UK.



The body hydraulic: Descartes' illustration of the sensory path of burning pain.

Emortality — at last!

Ali Zaman's recipe for eternal life will set us free.

Brian Stableford

It is still possible, of course, that we who are beneficiaries of the Zaman Transformation will not prove to be as long-lived as we dare to hope. Only time will tell whether those of us whose lives are not prematurely terminated by violence will live for 500 years, 5,000 years or five million. Pessimists are right to call attention to the fact that previous technologies of longevity have all failed to live up to the expectations of their inventors, although optimists are equally justified in pointing out that what we learn from our mistakes invariably helps us to avoid them, and that there was always bound to come a time when all our errors had been corrected.

It is arguable that none of the previous technologies of longevity were failures, even though they did not provide their beneficiaries with true emortality. Ironically, they were victims of their own success. The techniques of rejuvenation pioneered in the twenty-first century by Morgan Miller did indeed rejuvenate all the tissues and organs of a mammalian body, including the brain. Unfortunately, 'rejuvenating' the brain restored all the functionality of all the synaptic connections whose selective elimination is responsible for the formation of the individual personality and the development of rational intelligence. Even animals could not survive such an extreme restoration of innocence, although the boost that Miller's research gave to the production of tissue-culture foodstuffs allowed for the early release of most of the kindred species that humankind had enslaved for meat production.

The nanotechnological processes of organic repair developed in the wake of the Plague Wars by such megacorporations as PicoCon and OmicronA allowed the Miller Effect to be circumvented, but soon revealed the next hurdle facing architects of emortality. By the middle of the twenty-fourth century it was obvious that, although the nanomachines maintained existing synaptic connections, rather than restoring those whose extinction had been part and parcel of the process of individualization, they were too scrupulous in so doing. Eventually, their assiduous maintenance resulted in the 'robotization' of the mind, stifling its ability for further development and preventing the further evolution of personality. Although lesser animals equipped with the best repair nanotech could be reckoned emortal, humans could not.

Unfortunately, the apparent success of nanotechnological repair had resulted in an

excessive concentration of research efforts. The possibility of using genetic engineering to incorporate longevity into the human genome had been neglected for some time, and the kinds of transformation investigated by Ali Zaman might have been fully explicated a century earlier had it not been for overt and covert prejudice against that kind of approach to the problem. Once the necessity was clear, though, genetic engineers resumed the quest to provide a genetic blueprint for a brain capable of containing a mind that could continue to develop indefinitely without falling prey to the Scylla of the Miller Effect or the Charybdis of robotization. The advances in technique made in the interim by flower designers helped them to produce encouraging results without overmuch delay.

There remains a possibility that today's Natural Emortals will eventually reveal some other as-yet-unimagined vulnerability that will prevent them from extending their health and happiness indefinitely, but we are fully entitled to hope that the last major problem has been solved. The legless fabers who will presumably become the primary colonists of all low-gee environments within and beyond the Solar System will undoubtedly require more complicated Zaman Transformations, especially if they are to develop six- or eight-handed variants, but that is a matter of mere technical adjustment.

The champions of cyborgization are prob-

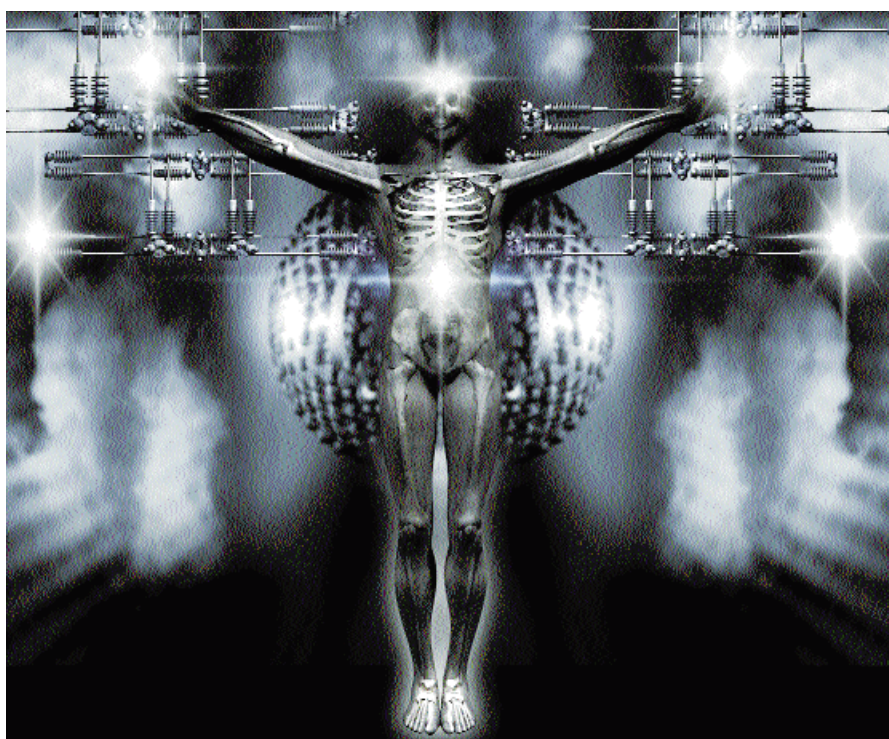
ably right to argue that the brains we have at present are inadequately pre-adapted for the kinds of neuromechanical augmentation that they are enthusiastic to investigate, but it is unlikely that biotechnological pre-adaptation will compromise the effects of the Zaman Transformation. There is, therefore, every reason to suppose that the various human-descended species that will undertake the boldest adventures in universal exploration will also share the gift of true emortality.

We shall never be entirely free of the shadow of death. There remain a dozen brutally simple ways in which human beings can be destroyed, no matter how clever their internal technology may be in compensating for injury or oxygen starvation. We have, however, relegated death to its proper place in human affairs: its threat may not be neutralized, but it is minimized. That is why I am now able to plan the compilation of a definitive history of death, whose conclusion will mark the most crucial boundary in human evolution.

Until now, all human history has been the history of death, but from this day forward all histories except mine will be the history of life. (From the introduction to *The History of Death* by Mortimer Gray, deposited in the Hypertextual Labyrinth in 2614.) ■

Brian Stableford's most recent novel is *Architects of Emortality (Tor)*, part of a future history that will be continued in 2000 in *The Fountains of Youth*.

e-mail: bstableford@cix.co.uk



OLIVER BURSTON

Plug-in quantum software

John Preskill

Some quantum states are hard to create and maintain, but are a valuable resource for computing. Twenty-first century entrepreneurs could make a fortune selling disposable quantum states.

In principle, quantum computers could solve problems that conventional digital computers could never solve¹. In practice, quantum hardware is in its infancy, and has far to go before quantum computers can realize their promise. Perhaps someday, as suggested on page 390 of this issue by Daniel Gottesman of Microsoft Research and Isaac Chuang of the IBM Almaden Research Center², the capabilities of quantum computers will be significantly extended by 'quantum software' that can be prepared offline and shipped to users over the 'quantum Internet'.

Whereas conventional computers process information encoded in bits, a quantum computer processes information encoded in quantum states, such as the internal electronic states of individual atoms, or the spin states of atomic nuclei. Existing quantum computers contain just a few atoms, and can perform only very simple computations. But because the inherent complexity of a quantum state rises very steeply as the number of atoms increases, a quantum computer acting on thousands of atoms could have staggering power. For example, today's digital supercomputers would take billions of years to find the prime factors of a number that is hundreds of digits long, whereas a quantum computer of the future might perform that task in seconds. Constructing large-scale quantum computers will be a formidable technological challenge, so it is not yet possible to predict when this revolutionary technology might become reality.

A quantum software program is a particular quantum state that enables a quantum computer to perform a specific task. If that quantum state is difficult or inconvenient to prepare, the user of a quantum computer might prefer to acquire the state from a vendor, rather than prepare it herself. The user's hardware then acts on the quantum state according to a standard protocol, but the outcome varies depending on the version of the software. Unfortunately for the user (but to the delight of the vendor) quantum software would be a consumable product, unavoidably damaged after a single use (Fig. 1). Thus we can foresee the flourishing of a quantum software industry. A manufacturer can design a valuable quantum state, and use a special-purpose device to churn out multiple copies of the state; these can be tested to assure quality and stored until needed. Con-

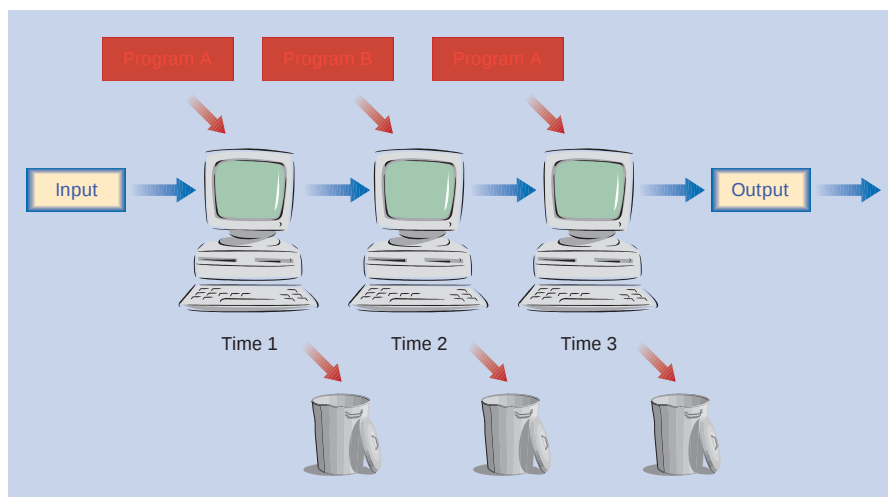


Figure 1 **Disposable quantum software.** Gottesman and Chuang² suggest that perhaps, one day, quantum software might be delivered from vendor to user over a quantum communication network (the 'quantum Internet'). After a single use, the software is irreparably damaged and must be discarded. To execute a quantum algorithm, the user might have to download and consume a particular program many times.

sumers can download the state for a fee and plug it into their own quantum computers to achieve improved performance.

One important application for quantum software will be to ensure that quantum computers function reliably. A quantum computer is equipped to execute a standard set of unitary transformations, called its quantum gates. By executing many gates in succession, a computer applies a complex unitary transformation to an input quantum state, yielding an output state that is ultimately measured. The set of standard gates may be finite, but a well-chosen gate set is universal, enabling the computer to simulate any specified unitary transformation to any desired accuracy³.

Because complex quantum states are extraordinarily fragile, the quantum-computing hardware that implements the gates needs to meet very demanding specifications. But even with superb hardware, a computer could achieve acceptable reliability only by applying the recently discovered principles of quantum-error correction: if quantum information is cleverly encoded, it can be protected both from the destructive effects of uncontrolled interactions with the environment, and from the cumulative effect of the inevitable small imperfections in the hardware^{4,5}. Furthermore, for a given coding scheme, particular quantum gates

that are well matched to the structure of the code can be executed fault tolerantly — that is, even though the hardware that implements the gate is imperfect, the action on the encoded quantum information is highly accurate. Fault-tolerant gates that form a universal set allow a quantum computation to be executed with good reliability.

For each of the known quantum-error correction schemes, some of the gates in the fault-tolerant universal set are easy to perform, whereas others are hard. These 'hard' gates would be most conveniently executed with quantum software that can be prepared ahead of time and then consumed during the operation of the gate. Realizing the gate with software rather than hardware is preferable because we could verify before use that the software has been prepared according to specifications, and we can discard or repair it if it is found to be faulty. In contrast, if our hardware fails badly during the execution of a quantum gate, it might be difficult to recover from the damage.

The idea of preparing special states offline to allow fault-tolerant quantum computation is not new^{6,7}. But Gottesman and Chuang² now propose a novel and more systematic approach to the preparation, use and classification of these states. Their central idea is that applying a quantum software routine can be viewed as a generalized form of

quantum teleportation. In standard teleportation⁸ one party (Alice) destroys an unknown quantum state, and then sends a classical message to another party (Bob), who proceeds to reconstruct a perfect replica of the original state. To achieve this feat, the two parties consume some standard quantum software that they share. Gottesman and Chuang note that if the software is suitably modified, the reconstructed quantum state is modified too. Instead of winding up with Alice's original state, Bob reconstructs a state to which a quantum gate has been applied. If Alice and Bob use different software, they execute a different gate. Standard teleportation has been demonstrated in the laboratory^{9–11}, and some of the protocols suggested by Gottesman and Chuang are feasible, or nearly so, with existing techniques.

A quantum computer with sufficiently sophisticated hardware could run on classical software. But a more affordable machine might be marketed with only rudimentary hardware tools. Unable to execute a complete universal set of fault-tolerant gates with its hardware alone, it would achieve that capability through quantum software suited to its coding scheme. Users of such computers would require broadband access to the quantum Internet¹², over which reliable quantum software could be shipped on demand.

Sometime during the twenty-first century (no one can say just when), fault-tolerant quantum computers made possible by quantum software may achieve processing speeds far surpassing those of conventional digital computers. But before that happens, quantum software could already be a marketable commodity. Because tampering with a quantum state leaves a detectable imprint, a shared quantum state can enable two parties to establish a secure communication link^{13,14}. To exploit the growing demand for privacy, a quantum software company might create a product that allows two customers to talk to one another without fear of eavesdropping, or to authenticate each other's identities. Although quantum computers will be needed to create and distribute the software, these devices would be less complex and easier to build than quantum computers that solve complex computational problems.

Another intriguing development is the recent discovery¹⁵ by Daniel Jonathan and Martin Plenio of Imperial College, London, that in a particular setting quantum software can serve as a catalyst for a quantum operation — it enables execution of the operation without being consumed in the process. Will quantum states be found that can catalyse fault-tolerant gates or other useful operations? The invention of reusable quantum software would be deflating news for the stockholders of 'Quantum Valley', but would be welcomed by tomorrow's consumers of quantum information technology. ■

John Preskill is in the Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, California 91125, USA.
e-mail: preskill@theory.caltech.edu

1. Shor, P. in *Proc. 35th Annu. Symp. on Foundations of Computer Science* (ed. Goldwasser, S.) 124–134 (IEEE Press, Los Alamitos, California, 1994).
2. Gottesman, D. & Chuang, I. L. *Nature* **402**, 390–393 (1999).
3. Deutsch, D. *Proc. R. Soc. Lond. A* **425**, 73–90 (1989).
4. Shor, P. *Phys. Rev. A* **52**, R2493–R2496 (1995).
5. Steane, A. M. *Phys. Rev. Lett.* **77**, 793–797 (1996).
6. Shor, P. in *Proc. 35th Annu. Symp. on Foundations of Computer Science* (ed. Goldwasser, S.) 56–65 (IEEE Press, Los Alamitos, California, 1994).

7. Knill, E., Laflamme, R. & Zurek, W. H. *Proc. R. Soc. Lond. A* **454**, 365–384 (1998).
8. Bennett, C. H. et al. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
9. Bouwmeester, D. et al. *Nature* **390**, 575–579 (1997).
10. Boschi, D. et al. *Phys. Rev. Lett.* **80**, 1121–1125 (1998).
11. Furusawa, A. et al. *Science* **282**, 706–709 (1998).
12. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).
13. Bennett, C. H. & Brassard, G. in *Proc. IEEE Int. Conf. Comp. Sys. Sig. Process.* 175 (IEEE Press, New York, 1984).
14. Ekert, A. K. *Phys. Rev. Lett.* **67**, 661–663 (1991).
15. Jonathan, D. & Plenio, M. B. *Phys. Rev. Lett.* **83**, 3566–3569 (1999).

Botany

The family tree flowers

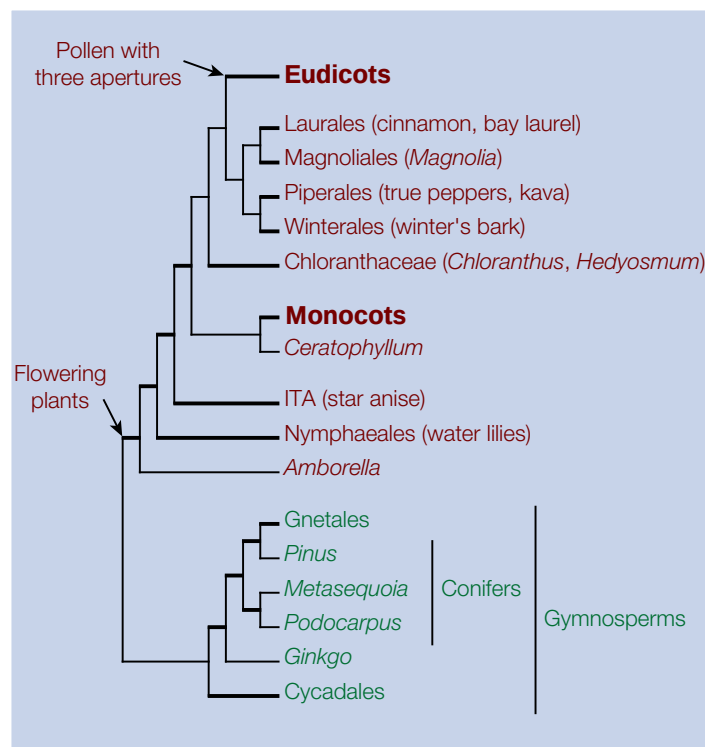
Paul Kenrick

The origin and relationships of flowering plants are among evolutionary biology's enduring mysteries. These comparative newcomers to the evolutionary stage number a staggering 250,000 living species classified into about 350 families. One of the cherished goals of botany is to unravel the complicated family tree, a process that involves establishing the branching relationships among species (phylogeny). Two papers in this issue take an impressive step forward. Soltis *et al.* (page 402)¹ and Qiu *et al.* (page 404)² use a multigene approach to identify the closest living relatives of flowering plants and map out the deepest branches of the family tree.

Analysing the deep phylogenetic rela-

tionships in flowering plants is a complicated business. This is partly because of the sheer size and diversity of the group, and partly because of its rapid diversification during the Early Cretaceous (130–90 million years ago)³. Flowering plants differ considerably from their closest living relatives in the gymnosperms (conifers and their allies), creating additional problems for 'rooting' the family tree. Recent molecular systematic work — which includes some of the largest phylogenetic analyses attempted for any group of organisms — confirms the existence of a major group (eudicots) which has a characteristic, three-aperture pollen type and includes most of the dicotyledons^{4,5}. The widely recognized division between

Figure 1 Summary of relationships among the major groups of flowering plants and gymnosperms based on the multigene sequence analysis of Qiu *et al.*². Thicker branches are those with strongest (90–100% bootstrap) support from the analysis. In this scheme, dicotyledons are not recognized as a group within the flowering plants.



dicotyledons and monocotyledons based principally on embryology (number of seed leaves), leaf morphology (veination type) and floral organization (number of parts) is not upheld. Furthermore, it is now clear that several of the traditionally recognized subclasses, such as the Dilleniidae and Hamamelidae, are not natural groups.

As to the earliest branches of the flowering-plant tree, there is strong evidence that these lie within the so-called magnoliid dicotyledons, a small group (only about 3% of living species diversity) which includes such familiar forms as water lilies (Nymphaeales), true peppers (Piperales), magnolias and several other related families such as the Winteraceae, Chloranthaceae and Lauraceae. But despite these and other successes, the exact pattern of branching among basal lineages has remained obscure.

One weakness of current approaches is that they are built mainly on single-gene data sets. Soltis *et al.* and Qiu *et al.* rectify things by combining gene sequences from different genomes (nucleus, mitochondrion, chloroplast) in individual species. As well as increasing the total data, this strategy is designed to reduce noise generated by gene-, function- and genome-specific phenomena. The results have broad implications for our understanding of flowering-plant evolution.

The two papers take slightly different approaches. Qiu *et al.*² amass a huge sequence database (five genes; 8,733 base pairs) for a comparatively modest number (105 species) of exemplar organisms; these are species selected to represent larger groups. Soltis *et al.*¹ sacrifice sequence length (three genes; 4,733 base pairs) for breadth of taxonomic coverage (560 species).

In their analysis, Qiu *et al.* provide what is arguably the first really solid evidence for the basal three branches of the flowering-plant tree (Fig. 1). They show that the family Amborellaceae, which consists of the single species *Amborella trichopoda*, is the sister group to all other flowering plants, a result that is consistent with another analysis based on duplicate phytochrome genes⁶. Next come the water lilies (Nymphaeales), followed by a group dubbed ITA (Illiciaceae, Trimeniaceae, Austrobaileaceae, Schisandraceae). It is the strength of evidence on these branches that really distinguishes this result from other large analyses based on fewer genes. It really looks as though Qiu *et al.* might have got the answer to the seemingly intractable problem of basal lineages in flowering plants.

How does the approach of Soltis *et al.* perform on this question? The tree topology is identical, but branch support for basal lineages is stronger in the study of Qiu *et al.*. This would seem to vindicate Qiu and colleagues' approach: at this level it is not so much adding more taxa that yields better

supported results but having more data from a comprehensive selection of exemplars. One must however be cautious with the exemplar approach, because it can lead to blind spots. *Amborella*, for example, did not appear in some earlier analyses because it was not a first-choice exemplar, and it is inconveniently confined to the Pacific island of New Caledonia!

The interest of the paper by Soltis *et al.* is in the breadth of taxonomic coverage. The general pattern of relationships is similar to that outlined by Qiu *et al.*, but there is more detail within the monocotyledons and eudicots. The relationships of model organisms and crops — such as *Arabidopsis thaliana* (thale cress), *Nicotiana tabacum* (tobacco), *Brassica oleracea* (cabbage, kale) and *Solanum tuberosum* (potato) — are defined more precisely. This information is essential for formulating general principles about the biology of flowering plants from particular experimental observations. Among other things, the analysis strongly supports a close relationship between families of nodule-forming nitrogen-fixing plants and the multiple origins of petals from sepals. Here we indeed have the beginnings of a general phylogenetic framework for flowering plants — a powerful research tool for comparative biology.

One surprising outcome of Qiu and colleagues' analysis is that it refutes the widely

held 'anthophyte hypothesis' for the origin of flowering plants. This hypothesis links flowering plants to a small and obscure group of living gymnosperms called Gnetales and the extinct Bennettitales, based on morphological similarities⁷. The molecular data suggest otherwise. Qiu *et al.* provide strong evidence that Gnetales group within the conifers to the exclusion of flowering plants. If this shift in relations holds up, it will have far-reaching consequences for our understanding of the evolution of such flowering-plant characteristics as the carpel, double fertilization and the flower itself^{7,8}. Furthermore, it implies that the split between the flowering-plant lineage and other living gymnosperms occurred much earlier than currently supposed, possibly by the end of the Carboniferous, 290 million years ago. ■

Paul Kenrick is in the Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

e-mail: P.Kenrick@nhm.ac.uk

1. Soltis, P. S., Soltis, D. E. & Chase, M. W. *Nature* **402**, 402–404 (1999).
2. Qiu, Y.-L. *et al.* *Nature* **402**, 404–407 (1999).
3. Friis, E. M., Pedersen, K. R. & Crane, P. R. *Ann. Missouri Bot. Gard.* **86**, 259–296 (1999).
4. Soltis, D. E. *et al.* *Ann. Missouri Bot. Gard.* **84**, 1–49 (1997).
5. Chase, M. W. *et al.* *Ann. Missouri Bot. Gard.* **80**, 528–581 (1993).
6. Mathews, S. & Donoghue, M. J. *Science* **286**, 947–950 (1999).
7. Doyle, J. A. *Int. J. Plant Sci.* **157**, S3–S39 (1996).
8. Friedman, W. E. & Carmichael, J. S. *Int. J. Plant Sci.* **157**(6 Suppl.), S77–S94 (1996).

Astronomy

Clouds from near and far

W. Butler Burton and Robert Braun

Scattered across the astronomical sky are hundreds of clouds of hydrogen gas, moving at such high velocities that they clearly stand out from the usual interstellar material in the Milky Way. These high-velocity clouds (HVCs) have remained enigmatic since their discovery in the 1960s¹. No stars associated with the gas have been detected; nor have infrared emission from warm dust or emission lines from simple molecules such as carbon monoxide, which might reveal a potential for star formation on a large scale. Direct distance measurements — on which knowledge of crucial physical parameters such as mass, density and size depend — are lacking.

Two papers in this issue, by Richter *et al.*² on page 386 and by Wakker *et al.*³ on page 388, report surprisingly different compositions for two nearby HVCs, and come to different conclusions about their origins. Many astronomers have attempted to explain the anomalously high velocities, and to place the HVCs in an appropriate astrophysical context, but no consensus has been reached⁴.

Perhaps no single explanation will suffice, as the work of Richter *et al.* and Wakker *et al.* appears to suggest.

Some objects listed as HVCs belong to a long stream of debris associated with the Large and Small Magellanic Clouds. The Magellanic Clouds are a pair of nearby dwarf galaxies that are gravitationally linked to the Milky Way, and which will eventually merge with our Galaxy. In its current orbit, the Large Cloud has just passed through the outer reaches of the Milky Way, and is now gravitationally dragging debris composed of interstellar material, as well as some stars, drawn from our own Galaxy. This tidal debris accounts for the high-velocity emission observed in the Magellanic Stream, but is not typical of HVCs in general.

Two competing explanations exist to account for the majority of the HVC emission that is not associated with the Magellanic Clouds. First, the Local Group hypothesis suggests that the HVCs are scattered throughout the Local Group of galaxies, at typical distances of some hundreds of kilo-

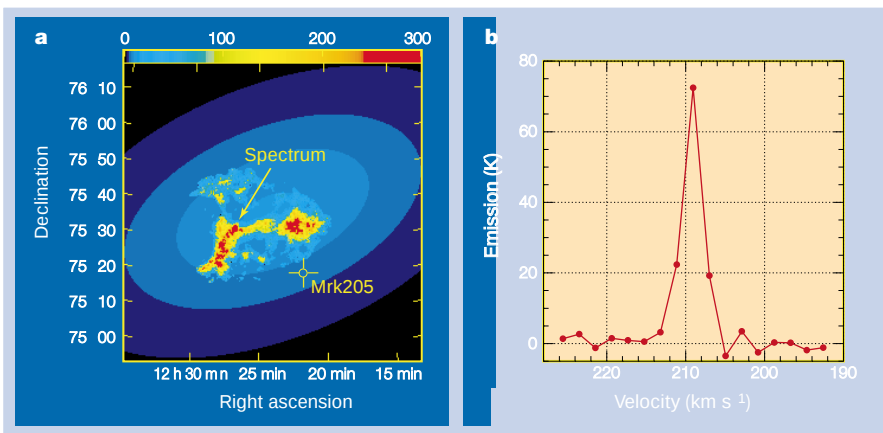


Figure 1 A compact high-velocity cloud (HVC) in which one or more cold cores are embedded in a diffuse, warmer halo. a, High-resolution image captured by the Westerbork radio interferometer. b, The hydrogen spectrum of one of the cores reveals a narrow line width — one of the narrowest H I emission lines ever observed — indicating a core temperature less than 85 K. This cloud is thought to be as much as 0.5 to 1 Mpc away. Larger complexes of HVCs, such as the one studied by Wakker *et al.*³, may be similar to the object shown here, but much nearer and in the process of merging with our Milky Way. Colour scale shows hydrogen flux in 10^{18} cm^{-2} .

parsecs (1 kpc = 3,262 light years). The possibility that HVCs are large gas clouds falling into the Milky Way has been discussed for more than 30 years, but has been rejuvenated⁵ by the idea that HVCs are the basic building blocks involved in the formation and growth of galaxies. In the Local Group hypothesis the gas would not be enriched in heavy elements produced by stellar evolution.

Second, there is the 'galactic fountain' model⁶, in which HVCs would be generated by supernova explosions in the Galactic disk. Heated gas would be ejected outwards into the Galactic halo where it would cool and then fall back towards the disk as HVCs. The fountain gas would be substantially enriched in heavy elements, because moderately high amounts in the Milky Way gas would be further enhanced by the supernova ejecta.

These alternative explanations for the origin of HVCs involve fundamentally different physics: in the Local Group hypothesis, the HVCs would be massive objects infalling from large distances, and bringing relatively unprocessed material into the Galaxy. In the galactic fountain model, the objects would be the result of violent events in our own Galaxy, and contain gas enriched by the processes of stellar birth and death. The competing explanations each provide a range of testable predictions. Measurements of the heavy-element abundance in HVCs could be crucial to distinguish between them.

Wakker *et al.*³ report the detection of ionized sulphur in an HVC complex seen in absorption against a distant galaxy. The amount of absorption indicates a heavy-element abundance of 9% of the solar value. This low metallicity (astronomers curiously refer to elements beyond helium as 'metals') is consistent with the infall of fairly pristine gas. The Wakker *et al.* result is consistent with

the Local Group hypothesis, even though it samples an HVC that is relatively nearby, perhaps 10 or 20 kpc away, in the outer halo of the Galaxy. In the Local Group hypothesis, such a complex would be a galactic building block that has come from far away to merge with the Milky Way. Indeed, observations of absorption lines against distant stars have placed a similar complex in the halo of the Milky Way⁷. Both these situations are reminiscent of the Sagittarius dwarf galaxy, which has ventured within the halo of our Galaxy, and is fuelling the growth of the Milky Way.

Yet Richter *et al.*² report the discovery of molecular hydrogen in another HVC (H_2 forms most efficiently on the surface of dust grains, which are composed of heavy elements) as well as a relatively high iron abundance, roughly half the solar value. The metallicity, and the constraint on distance (less than 55 kpc) that follows from having detected absorption against stars in the Magellanic Cloud, lead Richter *et al.* to favour a Galactic origin for the cloud in question and the fountain model to explain its high velocity. So from two measurements of heavy-element abundance we have two rather different conclusions.

To place these results in a more general context it is worthwhile considering more closely the two objects being studied. The Richter *et al.* measurement probes an extended region of high-velocity gas that varies between 70 and 140 km s^{-1} . Although the detailed distribution of this gas is still not clear, it lies in a region cluttered with debris from the Magellanic Clouds⁸. There is substantial uncertainty then, whether the object sampled by Richter *et al.* is a 'typical' HVC, from which far-reaching conclusions can be drawn.

What about the HVC complex probed by Wakker *et al.*? A large fraction of the total

emission due to high-velocity hydrogen gas can be related to the handful of large HVC complexes. On the other hand, most HVCs are compact and isolated objects, which can be distinguished from the complexes⁹. So again, a single nearby complex may not be representative of all HVCs. It is fair to say that compact HVCs would be ideal targets for the sort of absorption experiments reported in this issue, yet observing them is not easy. They are small, so the probability of finding a suitable background source is low. Plus they have substantial substructure, requiring high-resolution imaging.

Only a few HVCs have been targeted by radio interferometers producing high-resolution images^{10,11}. A compact HVC captured by radio interferometry reveals several cold cores embedded in a diffuse, warmer halo (Fig. 1). The diffuse halo of the object overlaps the Seyfert galaxy Mrk205, where magnesium absorption has been measured¹². The relative abundance of magnesium to hydrogen corresponds to around 5% of the value in the neighbourhood of the Sun, lending support to the Local Group model for compact HVCs. Evidence against the fountain model comes from one of the cores: the narrow width of the hydrogen emission spectrum indicates its temperature is less than 85 K. It is difficult to imagine that such cold and quiescent conditions would be found after a violent birth, such as in a supernova explosion.

Measurements of the metallicity of HVCs have so far proved inconclusive. But more of these difficult measurements will be needed to finally decide between the competing theories. Until then, who can say for certain whether the HVC phenomenon mainly involves primitive objects falling in from large distances and fuelling the evolution of the Galaxy, or whether it represents the returning fall-out thrown up by violent events in our own Galaxy? ■

W. Butler Burton is at the University of Leiden Observatory, Leiden 2300 RA, The Netherlands.
e-mail: burton@strw.leidenuniv.nl

Robert Braun is at the Netherlands Foundation for Research in Astronomy, Dwingeloo 7990 AA, The Netherlands.

e-mail: braun@nfra.nl

1. Muller, C. A., Oort, J. H. & Raimond, E. C. R. *Acad. Sci.* **257**, 1661 (1963).
2. Richter, P. *et al.* *Nature* **402**, 386–387 (1999).
3. Wakker, B. P. *et al.* *Nature* **402**, 388–390 (1999).
4. Wakker, B. P. & van Woerden, H. *Annu. Rev. Astron. Astrophys.* **35**, 217–266 (1997).
5. Blitz, L., Spergel, D. N., Teuben, P. J., Hartmann, D. & Burton, W. B. *Astrophys. J.* **514**, 818–843 (1999).
6. Bregman, J. N. *Astrophys. J.* **236**, 577–591 (1981).
7. van Woerden, H., Schwarz, U. J., Peletier, R. F., Wakker, B. P. & Kalberla, P. M. W. *Nature* **400**, 138–141 (1999).
8. Putman, M. E. & Gibson, B. K. *Proc. Astron. Soc. Aust.* **16**, 70–76 (1999).
9. Braun, R. & Burton, W. B. *Astron. Astrophys.* **341**, 437–450 (1999).
10. Braun, R. & Burton, W. B. *Astron. Astrophys.* (submitted).
11. Wakker, B. P. & Schwarz, U. *Astron. Astrophys.* **250**, 484–498 (1991).
12. Bowen, D. V. & Blades, J. C. *Astrophys. J.* **403**, L55–L58 (1993).

What do yeast proteins do?

Philip Hieter

To understand how a cell works we need to know the function of every gene in its genome. Genome sequencing provides a tremendous amount of information for the development of global (so-called functional-genomic) approaches towards this goal, complementing and enhancing the more traditional (single-gene) approaches. But a challenge to functional-genomic approaches is the development of efficient interfaces between the large data sets, the reagents generated and the scientists analysing specific biological processes.

On page 413 of this issue, Ross-MacDonald *et al.*¹ describe a new genome-wide strategy for the large-scale analysis of gene function in the budding yeast *Saccharomyces cerevisiae*. This impressive work provides functional information for roughly one-third of all genes in the yeast genome, providing — on an unprecedented scale — direct access to information, mutant strains and cloned reagents for subsequent analysis by the scientific community.

The yeast genome sequence was completed three years ago². Analysis of the DNA sequence predicted approximately 6,200 genes, only 40% of which had previously been identified by standard experimental methods. This left about 3,700 genes for which no functional information was known. For the past two decades, individual laboratories have been using 'traditional' methods to work out the functions of small sets of proteins (several to dozens) involved in specific biological processes³. The standard approach has been to combine classical genetics (isolation and phenotypic analysis of mutants) and biochemistry (analysis of protein activity *in vitro*) with recombinant genetics (manipulating cloned DNA segments and reintroducing them at homologous sites in the yeast genome). These methods allow the physiological outcome of gene mutations in the cell to be tied to the biochemical activities of the corresponding gene products in the test tube.

More recently, yeast has been a test bed for the development and application of functional-genomic technologies. There are several reasons for this. First, its genome sequence is relatively simple (six times smaller than that of the nematode worm, and 200 times smaller than the human one). Second, the biological relevance of the large data sets generated can be assessed relatively easily, allowing the validity and power of the emerging technologies to be evaluated. These technologies include genome-wide messenger RNA expression profiling (using

oligonucleotide chips⁴, serial analysis of gene expression (SAGE)⁵, green fluorescent protein (GFP) fusion proteins⁶ and DNA microarrays⁷), systematic analysis of protein–protein interactions⁸ (two-hybrid screens, mass spectrometry), and systematic gene knockouts⁹ (precise gene deletions that are molecularly bar-coded).

Ross-MacDonald *et al.*¹ now describe the scale-up and improvement of a method for manipulating yeast genes on a genome-wide scale¹⁰. They have generated a comprehensive collection of yeast strains that allows differential gene expression in different cell populations to be analysed, gene-disruption effects to be assessed, and the localization of proteins in the cell to be determined. Each of these mutant yeast strains was created by inserting a defined segment of DNA into a single gene. To do this, the authors constructed a 'mini' transposon, which inserts itself into a fragment of yeast genomic DNA cloned into a bacterial plasmid (a dispensable genetic element maintained independently of the genome in a bacterial cell). The authors then introduced this DNA into yeast cells, and selected for those cells that had taken it up at its normal location in the yeast genome. To check that the insertion had landed 'in frame' (meaning that the DNA sequence could still be read correctly), the authors also engineered a β -galactosidase reporter gene into the transposon. Cells that

contain this appear blue when grown under certain conditions.

Using high-throughput methods, Ross-MacDonald *et al.* introduced about 92,000 independent insertional mutations into the yeast genome. They found 11,000 strains that expressed the β -galactosidase reporter gene and, for half of these, determined where in the genome the insertions were. As expected, most insertions occurred in genes already predicted from the genome sequence (known as annotated open reading frames, or ORFs). These insertions affected about 2,000 different genes. But the authors also discovered over 300 previously unidentified genes (so-called non-annotated ORFs, or NORFs).

To assess the effects of the insertion mutations, Ross-MacDonald *et al.* next introduced more than 7,000 such insertions into a haploid yeast strain (which contains only one copy of each gene, so the effects of the mutations are not masked by an intact copy of the gene). The authors then looked for clues about gene function, and these data are now available on the web¹¹. Ross-MacDonald and colleagues could also partially excise the transposon. This left a small segment to which monoclonal antibodies were directed, allowing protein localization to be determined within the cell by immunofluorescence microscopy (Fig. 1). Of 1,340 strains tested in this way, 201 were localized to a specific site in the cell, and 214 showed cytoplasmic staining.

The transposon system will also allow patterns of gene expression to be assessed by comparing expression of the β -galactosidase reporter gene in different cell populations. For example, the authors found that 31 different genes were specifically induced during meiosis (formation of gametes), only six of which were previously known to be induced. However, only 17 of the 31 genes identified were independently found by measuring RNA levels using microarray hybridization¹². This supports the need for subsequent validation using independent means. The β -galactosidase reporter method may, in fact, be more sensitive for detecting gene expression at the protein level, because the read-out is a fusion protein rather than messenger RNA.

The functional information provided by Ross-MacDonald *et al.* gives us an accelerated starting point for further analysis. But what is the best way to unravel the function of the newly discovered genes? Will it be through the application of genome-wide, high-throughput methods and analysis of the large data sets? Or will it be by using more traditional single-gene methods? The answer lies in the union of the two approaches and in increased interactions between scientists working in different organisms, providing complementary experimental approaches. To work out the mechanistic details we will still need the traditional combination of

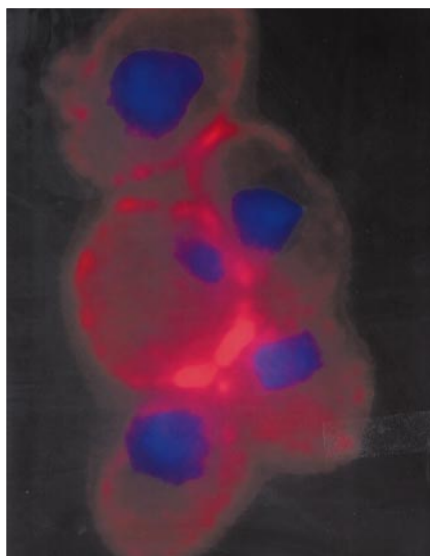


Figure 1 Protein localization in the cell. Ross-MacDonald *et al.*¹ probed cells with a monoclonal antibody against a fragment of the inserted 'mini' transposon (red). The cells were also stained with 4',6-diamidino-2-phenylindole (DAPI), which marks the DNA (blue).

genetics and biochemistry, focused on single genes and testable hypotheses. But functional genomics should make traditional approaches more efficient and productive, and will provide a new perspective on the behaviour of biological systems as a whole.

Philip Hieter is at the Centre for Molecular Medicine and Therapeutics, University of British Columbia, 950 West 28th Avenue, Vancouver, British Columbia, Canada V5Z 4H4.
e-mail: hieter@cmmmt.ubc.ca

1. Ross-MacDonald, P. *et al. Nature* **402**, 413–418 (1999).
2. Goffeau, A. *Science* **294**, 546–567 (1996).
3. Botstein, D. & Fink, G. R. *Science* **240**, 1439–1443 (1988).
4. Cho, R. J. *et al. Mol. Cell* **2**, 65–73 (1998).
5. Velculescu, V. *et al. Cell* **88**, 243–251 (1997).
6. Niedenthal, R. K. *Yeast* **2**, 773–786 (1996).
7. DeRisi, J., Iyer, V. R. & Brown, P. O. *Science* **278**, 680–686 (1997).
8. Fields, S. *Nature Genet.* **15**, 325–327 (1997).
9. Winzeler, E. *et al. Science* **285**, 901–906 (1999).
10. Burns, N. *et al. Genes Dev.* **8**, 1087–1105 (1994).
11. <http://ycmi.med.yale.edu/ygac/triples.htm>
12. Chu, S. *et al. Science* **282**, 699–705 (1998).

Ocean biogeochemistry

Sulphur in the mix

Ronald P. Kiene

The ocean surface layer influences climate not only through exchange of greenhouse gases such as carbon dioxide and methane with the atmosphere, but also gases like dimethylsulphide (DMS) which is thought to exert a cooling influence on the climate system. Climatic factors in turn influence the biogeochemical cycles of these trace gases, and there may then be feedbacks to the global climate system through the biota in the surface ocean. In 1987, Charlson *et al.*¹ proposed a phytoplankton–climate link that could operate through emission of DMS. Little was known about how climatic factors affect DMS production in surface waters, but Simó and Pedrós-Alió (page 396 of this issue²) now present evidence that the depth of sea-surface mixing, which largely depends on factors such as wind and temperature, may control DMS production and concentration.

Oceanic emissions of DMS are the largest biogenic source of sulphur to the atmosphere³, and are thought to have a cooling influence on global climate through formation of sunlight-scattering sulphur aerosols and cloud-condensation nuclei⁴. According to the feedback hypothesis of Charlson *et al.*¹, a global warming trend could be mitigated to some degree by an increase in DMS emission stimulated by the warming. Whether such a homeostatic feedback truly exists and how it might work has been difficult to ascertain, partly because both DMS and its phytoplankton-derived precursor, dimethylsulphoniopropionate (DMSP), have complex biogeochemical cycles in the surface ocean^{3,5}.

Simó and Pedrós-Alió² provide evidence that the short- and long-term response of the upper-ocean biogeochemistry to very shallow mixing conditions leads to increases in DMS concentration. Warmer climate conditions, leading to more intense or longer-lasting stratification (and hence shallower mixing), could therefore lead to more DMS emissions and a cooling influence on climate — as called for in Charlson and colleagues’

hypothesis¹. The strength of any cooling effect by increased DMS emission is still highly uncertain⁴, however, so the overall significance of the biogeochemical response to mixing remains to be determined. Moreover, any such feedback acting through the sulphur cycle must be considered together with other climatic effects of enhanced stratification. These effects could include lower uptake of carbon dioxide by the ocean⁶, leading to its accumulation in the atmosphere, which would act in opposition to DMS-induced cooling.

How might mixing depth control DMS accumulation in surface waters? An essential consideration here is that only a fraction of the phytoplankton DMSP undergoing degradation in sea water is converted to DMS. Based on their own data from the North Atlantic, and on data from the literature, Simó and Pedrós-Alió suggest that mixing-layer depth (driven by wind and temperature variations) controls the conversion efficiency of DMSP into DMS (hereafter the DMS yield) and ultimately the concentration of DMS in surface waters. They conclude that an unusual relationship exists whereby the DMS yield is highest (nearly 100%) in waters with a very shallow mixed layer, whereas it is lowest (about 10%) when mixing occurs to 10–20 m, and then higher again with deeper mixing to 30–65 m. At shallower mixing depths, plankton become more isolated from the rich nutrient reservoir below and are subject to more intense bombardment by solar radiation, including harmful ultraviolet-B.

The explanation given by Simó and Pedrós-Alió for the DMS yield pattern centres on the probable differential photo-inhibition of the phytoplankton (DMSP-producing) and bacterial (DMSP-consuming) populations that would be strongest, for the surface layer, under shallow mixing conditions. With few data available, there is still considerable uncertainty in the observed relationship, but the authors point out that



100 YEARS AGO

The scientific lessons of the war are crowding upon us. We have already referred to the blunder made by our military authorities in not sending Marconi apparatus to South Africa amongst the first equipments. We now learn indeed, after the investment of Ladysmith is drawing to a close, that Marconi apparatus is being sent out. The silence of Ladysmith during the last eventful weeks will point the moral, which is not likely to be forgotten in the future ... Some time ago the importance of a locomotive search-light in operations of war was strongly represented to the military authorities; but they would have none of it. Fortunately, however, the naval force in Natal has now provided the army with one.

As a proof of his cordial sympathy with the cause of bird protection, the Poet Laureate, Mr Alfred Austin, has written a special poem for the Christmas card which the Society for the Protection of Birds is issuing this year. It is entitled “Peace and Goodwill to the Birds,” and is illustrated by a coloured picture of that much persecuted bird the tern, designed for the purpose by Mr. A. Thorburn.

From *Nature* 23 November 1899.

50 YEARS AGO

Prof. H. Yukawa, who has been awarded the Nobel Prize for Physics for 1949, is best known for his theory of nuclear forces which, in 1935, first postulated the existence of a particle a few hundred times heavier than the electron. The nuclear forces would then bear the same relation to the possible emission and absorption of such a particle as the electromagnetic forces in an atom bear to the emission and absorption of light. The discovery of the meson in cosmic rays appeared to be a confirmation of Yukawa’s prediction, but the study of its properties gradually led to the conviction that it could not be identical with the particle required for Yukawa’s theory. It was not until 1947 that Powell and his collaborators demonstrated the existence of a second short-lived particle, the π -meson, which is known to be the parent of the cosmic-ray meson, and which is strongly linked to protons and neutrons. This provided a brilliant vindication of Yukawa’s idea. The detailed theory of the relation between this particle and the nuclear forces is still in its infancy; but, whatever the outcome, all thought about nuclear forces ... is entirely dominated by the ideas of Yukawa.

From *Nature* 26 November 1949.

the unusual pattern is similar to the effect of mixing depth on water-column-integrated photosynthesis in the Antarctic⁷. One key to interpreting the mixing-depth pattern might be that ultraviolet radiation or other factors such as nutrient limitation or grazing could diminish the assimilatory metabolism of DMSP⁵, thereby allowing more DMSP to be converted to DMS.

Regardless of the actual cause, the relationship between higher DMS production and shallow mixing observed by Simó and Pedrós-Alió is consistent with the ocean-wide build-up of surface DMS in summer, when stratification is greatest and mixing depths shallowest (see their Fig. 2 on page 397). What is especially intriguing about the 'summer DMS paradox' is that maximum DMS concentrations in surface waters are found at a time of year when phytoplankton biomass (as judged by levels of chlorophyll *a*) and DMSP concentration are well below their annual maxima.

The clear illustration of this discordance is a notable contribution because most researchers are still looking for close ties between DMS and algal biomass or DMSP. The data synthesis by Simó and Pedrós-Alió suggests that the crucial ties may lie elsewhere, perhaps in the food-web structure (for instance a shift to DMSP-rich picophytoplankton) or in the photophysical effects on DMS biogeochemistry. Processes such as photooxidation of DMS and the influence of algal-derived photosensitizers are

probably important^{8,9}. Inhibition of DMS-consuming bacteria by increased ultraviolet exposure in shallow mixing layers could also contribute to DMS build-up. Further research on what controls net production of DMS is necessary, and the mixing regime needs to be considered in any such work.

The studies of Simó and Pedrós-Alió and of others^{7,10} show that mixing of the upper ocean can have significant effects on biological and biogeochemical processes. These investigations indicate that the responses of the upper-ocean biogeochemical system to variations in climate-driven mixing are both complex and surprising. Clearly, we need a better understanding of how the upper ocean will respond in the context of events such as El Niño and future global warming. ■

Ronald P. Kiene is in the Department of Marine Sciences, University of South Alabama, LSCB-25, Mobile, Alabama 36688, USA.

e-mail: rkieni@jaguar1.usouthal.edu

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Simó, R. & Pedrós-Alió, C. *Nature* **402**, 396–399 (1999).
3. Andreae, M. O. *Mar. Chem.* **30**, 1–29 (1990).
4. Andreae, M. O. & Crutzen, P. J. *Science* **276**, 1052–1058 (1997).
5. Kiene, R. P., Linn, L. J., Gonzalez, J., Moran, M. A. & Bruton, J. A. *Appl. Environ. Microbiol.* **65**, 4549–4558 (1999).
6. Sarmiento, J. L., Hughes, T. M. C., Stouffer, R. J. & Manabe, S. *Nature* **393**, 245–249 (1998).
7. Neale, P. J., Davis, R. F. & Cullen, J. J. *Nature* **392**, 585–589 (1998).
8. Kieber, D. J., Jiao, J., Kiene, R. P. & Bates, T. S. *J. Geophys. Res.* **101**, 3715–3722 (1996).
9. Dacey, J. W. H., Howse, F. A., Michaels, A. F. & Wakeham, S. G. *Deep-Sea Res.* **45**, 2085–2104 (1998).
10. Karl, D. M. *et al.* *Nature* **373**, 230–234 (1995).

Phylogenomics

Ancestral primate viewed

Stephen J. O'Brien and Roscoe Stanyon

How did the genomes of modern mammals come to be organized the way they are? Improvements in genome sequencing, gene mapping and chromosome recognition mean that, by comparing the order and sequence of genes in different genomes, we may be able to trace the evolutionary pathways that our ancestors took. The potential of such an approach is illustrated by three new studies^{1–3}, which compare the genome maps and chromosomes of 15 species of primate, using four non-primate orders — carnivores, rodents, artiodactyls (hoofed mammals) and tree shrews — as a reference standard for the organization of ancestral mammalian genomes. Although the approach itself is not new, the chance to explore entire vertebrate genomes with such a high level of precision is.

Comparative cytogenetics (the study of differences in chromosome structure and appearance) dates back three decades. Then, factors such as the position of the cen-

tromere, a chromosomal structure involved in nuclear division, alerted researchers to differences between distantly related species⁴. Later, specific banding patterns (obtained by staining chromosomes) were used to infer sub-chromosomal homologies, and many of these have turned out to be correct⁵. Today, comparative gene maps have been started for over 40 mammalian species⁶. Chromosome painting (where flow-sorted individual chromosomes of one species are labelled with fluorescent dyes and hybridized *in situ* to chromosomes of another species) has allowed long stretches of conserved gene segments to be identified among very distantly related species.

Such comparisons have given us a glimpse of the genome rearrangements that punctuate adaptation and species formation⁷. Because exchanges between genomes are rare, and the junctions between gene segments can be easily identified, such rearrangements are powerful evolutionary characters that proscribe the

history of species that retain them. Comparative 'phylogenomic' approaches therefore combine what we know about segment homology with the strategies of evolutionary cladistics (a method of classification based on those shared characteristics that are assumed to indicate a common ancestry). In the first of the new papers¹, Haig restates an idea often heard from palaeontologists working with a tooth or cranial fragment, when he writes that describing the process of reconstructing ancestral chromosome structures "can be likened to the difficulty of writing an entertaining account of the piece-by-piece assembly of a jigsaw puzzle". The process is indeed tedious and treacherous, but the goal of reconstructing genomic history is a lofty one.

So how is it done? First, work out how many chromosomes or segments are conserved between the genomes of two species. Second, from this count, estimate how many chromosome exchanges (translocations, transpositions, fusions, fissions or inversions) it takes to rearrange one genome to the other — that is, how many scissor cuts would allow a human genome arrangement to be turned into that of a tree shrew, and vice versa? Next align additional species to the first two and identify the conserved homology blocks. The ancestral forms can be regarded as those shared between several species. (By contrast, derived segments or connections are unique to one or more closely related species, and serve as evolutionary signatures of particular lineages.) Finally, assemble the hypothetical ancestral genome — for example, for all primates — by comparison with the genomes of 'outgroup' species. In the new studies these included tree shrews; carnivores (cats, mink and seals); perissodactyls (horses); artiodactyls (cows, sheep and goats); or rodents (rats and mice). The ancestral chromosomal traits were defined as those shared between primates and the outgroups.

This process allows the minimum-sized fragments of chromosome exchange to be identified (termed the smallest conserved evolutionary unit segment; SCEUS⁸). The SCEUS is then treated as an evolutionary character. Haig¹ introduces a 'Cambridge grid', which tabulates the chromosomes of one species along the horizontal axis and those of a second species vertically. He then assigns individual Greek letters in the matrix to ancestral SCEUS associations. Once the ancestral genome is designed, the genomic exchanges that have occurred in lineages leading to the living species can be recapitulated.

The new studies^{1–3} interpret available data to reconstruct chromosomal characteristics from the ancestor of all primates (Fig. 1, overleaf). This creature last lived at least 65 million years ago, so it seems remarkable that, compared to the ancestral types, 18–20 human chromosomes remain unchanged and the rest have but a single exchange. Different exchanges have occurred in the lineages lead-

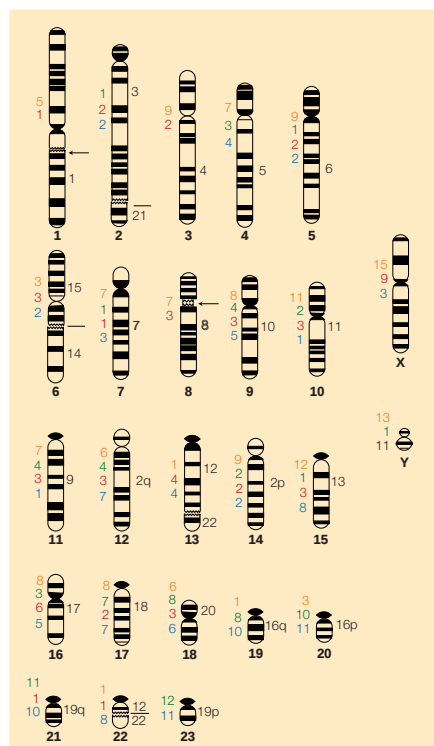


Figure 1 Reconstruction of the ancestral genome of living primates, depicted as conserved and rearranged human chromosomes¹⁻³. To the right of each ancestral primate chromosome is the human chromosome number. To the left are: the number of primate species (out of 15) in which the chromosome synteny is conserved as it occurs in the ancestral primate genome (yellow); the number, out of 15, in which a whole chromosome homology block is retained, but attached to another chromosomal piece (green); and the same counts for 11 outgroup species including pig, cow, muntjac, dolphin, cat, seal, mink, horse and shrew¹⁻³ (red and blue). Arrows indicate proposed separations of human chromosomes postulated by Müller *et al.*³ (chromosome 1) and Haig¹ (chromosome 8), which are unique to these authors.

ing to distinct primate families and genera. But, for most species, fewer than 20 rearrangements are enough to re-assort modern genomes to that of the primate ancestor. At least 12 of the ancestral primate chromosomes are also found intact in the genomes of humans, cats, seals, cows and shrews. So the ancestral primate genome must differ only marginally from the ancestral mammalian genome, which pre-dates the origins of carnivores, primates and artiodactyls.

The new papers also affirm that neither the positions of chromosomal breaks nor the evolutionary rate of rearrangements in primates seem to follow predictable patterns. Breaks can occur in many positions on any chromosome. Many species of primate, such as humans, great apes and Old World monkeys, show a remarkably slow rate of genome exchange — on the order of one or two exchanges every ten million years. By con-

trast, other species (gibbons, owl monkeys and lemurs, for example) globally reorganize their genomes, with two to four times more rearrangement relative to organization of the ancestral genome. A similar rate dichotomy is apparent in other mammalian orders, where a slow/default rate seen in most lineages (such as cat, mink, shrew and pig) is punctuated by drastic genome shuffles in others (dogs, bears, rodents and bears)⁷.

But there are limitations to what we can infer from comparative gene maps and chromosome painting. First, only a handful of mammalian species (human, mouse, rat, pig, goat, sheep, cow, cat and dog) have comparative maps with enough genes to allow their genomes to be compared^{6,7}. Second, although chromosome painting (on which most of the primate analyses are based) allows whole genomes to be assessed, small segments can be overlooked because of weak DNA hybridization. Precision is improved with reciprocal painting⁹⁻¹¹, or when more closely related index species are used as probes^{10,11}. Third, the order of the genes within homology segments was not considered in the primate assessment shown in Fig. 1. This means that an important class of chromosome exchanges, the interstitial inversions and translocations, is excluded. Comparative alignments of human gene order with ordered gene maps of non-primate species (goat and cat)^{12,13} have uncovered two to three times more breaks than were revealed by chromosome painting.

Another hurdle is analytical — there is no consensus algorithm for tracking genomic exchanges. Assumptions about the randomness of genome exchanges are probably oversimplifications, and there may be favoured site changes that introduce confounding convergent changes to evolutionary interpretations. So far, little emphasis has been placed on discordant exchange rates among

lineages, and for the different categories of exchange observed. Weighting of such characters in evolutionary analyses is an important consideration that needs theoretical and empirical input. Finally, the choice of species sampled could severely bias interpretations. Discrimination of shared-derived, as opposed to shared-ancestral arrangements, depends on how frequently they occur in the outgroups studied (Fig. 1). More is better in statistical terms, yet only a few species^{6,7,9} in fewer than half of the mammalian orders have been assessed.

Powerful new genomic and gene-mapping technologies should overcome these limitations. The resolution of comparative genome mapping is approaching its highest power ever, allowing linear maps of index species to be aligned explicitly. The primates offer the first and most accurate look at the history of human genome organization, but certainly not the last. We'll soon have similar reconstruction among other mammalian orders⁷ and even beyond, as previewed on page 411 of this issue¹⁴, where the chicken gene map weighs in on a comparative genomics perspective.

Stephen J. O'Brien and Roscoe Stanyon are in the Laboratory of Genomic Diversity, National Cancer Institute, Frederick, Maryland 21702-1201, USA. e-mail: obrien@ncifcrf.gov

- Haig, D. *Phil. Trans. R. Soc. Lond. B* **354**, 1447-1470 (1999).
- Chowdhary, B. P. *et al. Genet. Res.* **8**, 577-589 (1999).
- Müller, S. *et al. Chromosoma* **108**, 393-400 (1999).
- Hsu, T. C. & Bernirschke, K. (eds) *An Atlas of Mammalian Chromosomes* Folio 122 1774 (Springer, Berlin, 1969).
- Yunis, J. J. *et al. Science* **208**, 1145-1148 (1980).
- Graves, J. A. M. & Vandenberg, J. *Inst. Lab. Anim. Res. J.* **39**, 47-260 (1999).
- O'Brien, S. J. *et al. Science* **286**, 458-481 (1999).
- O'Brien, S. J. *et al. Nature Genet.* **3**, 103-112 (1993).
- Wienberg, J. & Stanyon, R. *Curr. Opin. Genet. Dev.* **7**, 784-791 (1997).
- Nash, W. *et al. Cytogenet. Cell Genet.* **83**, 182-192 (1998).
- Stanyon, R. *Cytogenet. Cell Genet.* **84**, 150-155 (1999).
- Schibler, L. *et al. Genome Res.* **8**, 901-915 (1998).
- Murphy, W. *et al. Genome Res.* (submitted).
- Burt, D. W. *et al. Nature* **402**, 411-413 (1999).

Ecology

Competition and coexistence

Ulrich Sommer

Early experimental and modelling work in community ecology¹ led to the principle of competitive exclusion. This states that, among two or more species of primary producer — that is, plants — competing for a shared resource, only the best competitor will survive². The apparent contradiction between competitive exclusion and the species richness found in nature has been a long-standing enigma. Explanations of why strong competitors do not inevitably have a stranglehold have invoked environmental disturbance and the differing resource requirements of different species.

On page 407 of this issue, however, Huisman and Weissing³ present a way to reconcile the persistence of diversity with competitive exclusion in undisturbed ecosystems with few limiting resources.

The contradiction between the principle of competitive exclusion and the diversity of natural communities first became obvious for phytoplankton (Hutchinson's 'paradox of the plankton'⁴). Planktonic algae share the same potentially limiting resources (light, carbon dioxide and mineral nutrients, namely nitrogen, phosphorus, silicon, iron and possibly a few other trace elements), are

suspended in a relatively well-mixed environment, and retrieve their nutrients from a common pool (the dissolved phase in the illuminated layer of water at the surface of seas or lakes). Experimental ecologists rose to Hutchinson's challenge, and their microcosm studies with planktonic algae became a cornerstone in developing both experiment and theory.

The contradiction between competitive exclusion and species richness appeared less severe for higher plants, because they are more segregated in space — for example, deep-rooted and shallow-rooted species might draw the same nutrient from separate pools. Nevertheless, controlled competition experiments with higher plants have also shown strong tendencies towards competitive exclusion.

Tentative explanations of the seemingly paradoxical diversity of plant communities have invoked the absence of competition in natural ecosystems; temporal and spatial variability of environmental conditions; and coexistence based on different optimal ratios of limiting resources. The article by Huisman and Weissing³ is a theoretical milestone merging the approaches based on variability and resource ratio.

The idea that diversity arises from spatial and temporal variability in the environment is based on the fact that competitive exclusion takes time, and that inferior competitors might find refuges in time or space. Such refuges could be sites or periods without competition or with reversed competitive hierarchies. The explanation has been formalized as the 'intermediate disturbance hypothesis'⁵ (IDH), which predicts a peak in species richness at intermediate intensities and frequencies of disturbance, thought of as events leading to environmental variability.

Culture^{6,7} and field⁸ experiments with phytoplankton confirmed the IDH, showing greatest diversity at disturbance intervals of three to six days, equivalent to up to ten generation times of unicellular algae. Comparative analyses of field data also supported the hypothesis⁹, although it proved difficult to quantify disturbance or environmental variability when it had not been imposed by the experimentalist, but rather consisted of changes in many potentially important controlling factors. All IDH-inspired studies have viewed disturbances as external to the community in question — for instance in the deepening of the surface-mixed layer of lakes and seas by wind and convective cooling. It was the implicit assumption that, with a constant supply of resources and constant physical conditions, primary-producer communities would approach a steady state, whereas oscillations or even chaotic dynamics were well known from models with several trophic levels.

The 'resource ratio hypothesis'¹⁰ (RRH) took a quite different approach. Primary

producers need essentially the same resources: light, carbon dioxide and mineral nutrients. However, they need them in different ratios. So resource supply might be balanced in a way that different species are limited by different resources and, thereby, coexist in perfect equilibrium at constant population densities. This was first shown by the coexistence of two diatoms with different optimal ratios of silicate and phosphate¹¹.

Further experiments with other combinations of algal strains from cultures or with naturally mixed assemblages of phytoplankton confirmed the stable two-species equilibrium based on two limiting resources. By extrapolation, it was thought that the same would happen if resource supply were balanced so that more than two resources were limiting — there was a commonly held belief that the number of coexisting species would equal the number of limiting resources. But this assumption was not tested experimentally or by modelling. Obviously, the RRH could not account for the full species richness of primary producers, because only a few resources (usually fewer than five) can become limiting in natural ecosystems.

This is the point where Huisman and Weissing³ step in. They show, by numerical modelling, that the competition dynamics of systems with more than two limiting resources are fundamentally different from those with only two limiting resources. Within a wide range of parameter values, sustained oscillations or even chaotic dynamics of resource concentrations and of species' abundances are possible, even under constant resource supply and constant physical conditions. These oscillations or chaotic changes create the environmental variability needed for the persistence of more species than the limiting resources would seem to allow.

This does not imply that external disturbances are unimportant in nature. Everybody knows that they occur and research has shown that they can promote diversity. But Huisman and Weissing's model has shown that externally undisturbed plant communities can produce their own disturbances. Now it is up to the experimentalists to catch up with the modellers. ■

Ulrich Sommer is at the Institut für Meereskunde, Düsternbrooker Weg 20, D-24105 Kiel, Germany.

e-mail: usommer@ifm.uni-kiel.de

1. Gause, G. J. *The Struggle for Existence* (Williams & Wilkins, Baltimore, 1934).
2. Hardin, G. *Science* **131**, 1292–1297 (1960).
3. Huisman, J. & Weissing, F. J. *Nature* **402**, 407–410 (1999).
4. Hutchinson, G. E. *Am. Nat.* **95**, 137–145 (1961).
5. Connell, J. H. *Science* **199**, 1302–1310 (1978).
6. Gaedeke, A. & Sommer, U. *Oecologia* **71**, 25–28 (1986).
7. Sommer, U. *Limnol. Oceanogr.* **40**, 1271–1277 (1995).
8. Flöder, S. & Sommer, U. *Limnol. Oceanogr.* **44**, 1114–1119 (1999).
9. Padisak, J., Reynolds, C. S. & Sommer, U. (eds) *Intermediate Disturbance Hypothesis in Phytoplankton Ecology* (Kluwer, Amsterdam, 1993).
10. Tilman, D. *Resource Competition and Community Structure* (Princeton Univ. Press, 1992).
11. Tilman, D. *Oecologia* **58**, 338–348 (1977).

Daedalus

Internal ecology

Last week Daedalus pointed out that a fetus in the womb will accept foreign tissue as part of itself, and will continue to accept that tissue in adulthood. So, he says, the way is open to creating mixed species — centaurs, griffins, sphinxes, unicorns and so on. To make a unicorn, for example, add a few rhinoceros cells to a horse fetus. The adult horse will then accept a rhino horn transplant. But surgery would probably work better on the fetuses. Fetal rhinoceros horn-bud tissue would be grafted onto the forehead of the fetal horse *in utero*. The growing fetus, with its greater capacity for repair and development, would probably solve the various compatibility problems better than a team of surgeons trying to marry up the inflexible adult tissues.

Biologists, not to mention owners of zoos, leisure parks and nature reserves, would be fascinated by the resulting chimerae. Even the classical chimera — part lion, part goat, and part snake — might be feasible. It would probably not breed true, or even at all, but would still be a biological triumph. Yet Daedalus has a more challenging goal — animal-plant chimerae. Imagine, he says, a combined man and green plant. The plant would photosynthesize, taking in carbon dioxide and producing oxygen and glucose; the man would conduct the reverse reaction. He would also produce urea and other metabolites useful to the plant — which would thus act as an extra liver and kidney. Such a chimera would be a most efficient self-contained ecosystem.

The problems are formidable. Even with immunocompatibility guaranteed, few plants could be genetically engineered to have a sap compatible with, or replaceable by, human blood. Even Daedalus is unlikely to create a man with real cauliflower ears. But a man with algae growing happily in his skin should be far more feasible. He would absorb sunlight, and feed, breathe and excrete internally, at least to a useful degree. The first 'little green men' will worry the flying-saucer cultists, and could arouse novel colour prejudice. But, with their wonderful ecological economy, they may be the way forward for humanity. Besides, with a little cunning plant-metabolic transfer, they could enjoy a constant internal supply of nicotine, caffeine, cocaine or cannabinoids. **David Jones**

The further Inventions of Daedalus (Oxford University Press), 148 past Daedalus columns expanded and illustrated, is now on sale. Special *Nature* offer: m.curtis@nature.com

Obituary

Charles M. Steinberg (1932–99)

On 17 September, Charley Steinberg died in Basel, Switzerland, of leukaemia that developed after a long-drawn-out illness. His name is not in the usual list of Heroes of the Molecular Biological Revolution, but to those who were fortunate enough to have come under his spell he was perhaps the supreme master of them all. Indeed, it is a sign of his special kind of mastery that he does not figure in the list.

Born in Montgomery, Alabama, Steinberg graduated from Vanderbilt University, Tennessee, in 1954. He was then drawn to Caltech by Max Delbrück (who, having himself worked at Vanderbilt during the Second World War, used to claim that Steinberg was the best thing ever to have come out of that university). During the next few years, the biology department at Caltech was the home for an astonishing array of graduate students and postdoctoral fellows — Edgar and Epstein, Streisinger, Meselson and Stahl, Drake, Rubin and Temin, and many others, all of whom were to make major contributions to the foundations of molecular genetics. But each would surely have agreed that Steinberg was the intellectual giant in their midst. Indeed when the physicist Richard Feynman decided that, for a while, he would try his hand at biology, he chose Steinberg (still a graduate student) to be his tutor and supervisor.

During those years at Caltech, Steinberg played a central role in two important acts of clarification. Certain bacteriophages had been shown to undergo genetic recombination when multiplying in their host. At that time nothing was known about the relation between DNA replication and recombination, and much effort therefore went into trying to unravel the mysteries of the phage-infected cell. Finally, in 1958 Steinberg and Frank Stahl produced a precise and elegant analysis that showed that phage genetics could not, in principle, show what kinds of interaction were going on between replicating DNA molecules, and this put a stop to much fruitless speculation.

Steinberg's other contribution was in his part in identifying and characterizing the different kinds of conditional lethal mutations in phage T4. The study of these mutations was crucial in elucidating the relation between DNA sequence, the genetic code and the nature of genes. It was during this collective work by Delbrück's students that Steinberg told Harris Bernstein that, if he would help count

culture plates, they would name any discovery after Bernstein's mother. This is why one of the three (suppressible) nonsense codons is now called amber (*Bernstein* in German).

After Steinberg completed his PhD thesis, Delbrück took him in 1961 to Cologne to help in the attempt to import the non-hierarchical American style to a German university. After two years, Delbrück returned to Caltech and Steinberg moved to the biology division at Oak Ridge, Tennessee. While there he was able to show the local experts (who had failed in their efforts) how to prepare the other radioactive isotope of phosphorus, ³³P, that was needed to resolve certain questions about the mechanism of DNA inactivation by ³²P and ³H decay. Later, it was the availability of two distinguishable radioisotopes of phosphorus that allowed Hamilton Smith to put different labels into the two strands of DNA and determine the mechanism of action of the class II restriction enzymes that were to prove to be nature's gift to sequencers.

In 1970, Niels Jerne became head of the new Hofmann-LaRoche Institute of Immunology in Basel. He immediately recruited Steinberg to be one of its first tenured members of staff and to serve as intellectual watchdog for the institute (to place some restriction on Steinberg's critical instincts, Jerne also put him in charge of complaints). During the next 20 years, the institute must have exceeded its creators' fondest hopes, for it was largely responsible for showing that antibody diversity is generated first by the shuffling of segments of genes to give multiple permutations and combinations, and second by high levels of random mutation selectively directed to what are called the hypervariable regions of antibodies.

Steinberg played a crucial part in each of these discoveries. Unfortunately, that was not obvious from a casual reading of the literature because, after doing much of the work, he usually chose to retreat into his role as unseen

A self-effacing influence in genetics and immunology

advisor. When first Jerne and then Susumu Tonegawa were awarded Nobel prizes for their work on antibody formation, many felt that Steinberg could well have shared in one or other of the prizes. His contribution is there for all to see, however, in the clarity of the annual reports of the institute. Until 1995, Steinberg wrote a dazzling introduction to each report and edited each laboratory's contribution. If you want to observe a fine scientific mind at work, read those reports.

Once upon a time, Steinberg's erstwhile student, The Famous Physicist, was talking to the director of the Oak Ridge National Laboratory. Feynman remarked that he knew a clever fellow at Oak Ridge named Charley Steinberg, who was the smartest guy he had ever met. When the director got back to Oak Ridge he asked about this Steinberg person and learned that he was one of the more influential scientists in the biology division. But Steinberg never put himself forward because he despised self-promotion in others.

To young scientists, who have to work in a world where survival depends on continued visibility, such a self-effacing life devoted to precision and excellence must seem scarcely believable. Yet, to all those friends and associates who constantly sought his advice, Steinberg represented the ideal combination of brilliance and integrity. For him nothing was difficult. And perhaps for that reason he chose to give credit to everyone but himself.

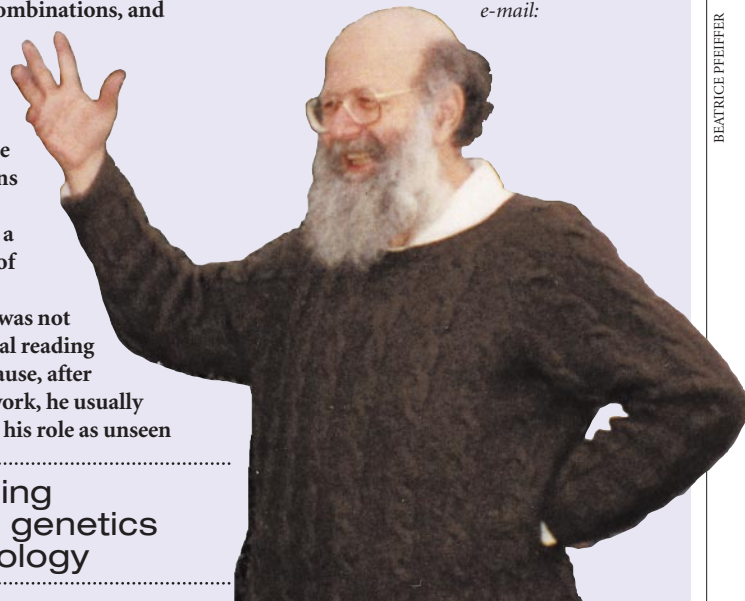
Jack von Borstel and John Cairns

Jack von Borstel is in the Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9.

e-mail: vborstel@ualberta.ca

John Cairns, formerly of the Harvard School of Public Health, Boston, is at Hollygrove House, Wilcote, Charlbury, Oxfordshire OX7 3EA, UK.

e-mail:



BEATRICE PFEIFFER

Feeding by mandibular raking in a snake

The tiny threadsnake has a unique way of devouring ants before they can strike back.

Most snakes transport prey through the mouth by using asynchronous ratcheting movements of their upper jaws^{1,2}. In contrast, we have found that threadsnakes (members of the basal snake clade Scolecophidia) have a unique feeding mechanism in which the tooth-bearing elements of the lower jaw rotate synchronously in and out of the mouth, dragging prey into the oesophagus. This mechanism, which we call 'mandibular raking', is the only vertebrate feeding mechanism known in which prey is transported exclusively by movements of the lower jaw.

Threadsnakes (family Leptotyphlopidae) are tiny, burrowing serpents that feed predominantly on the larvae, pupae and adults of social insects³ (Fig. 1a). They are rarely encountered in the wild and, because of their specialized diet, are difficult to maintain in captivity. As a result, little is known about their natural history or about how they feed^{4,5}: their extremely small size and ventrally placed mouths (Fig. 1a–c) also make it difficult to study their feeding mechanics.

We used an inverted dissecting microscope coupled to a high-speed video system to study ingestion and prey transport in the threadsnake *Leptotyphlops dulcis*. The imaging apparatus was positioned beneath a clear Plexiglas feeding chamber to record an unobstructed, highly magnified view of the snake's subterminal mouth.

This technique revealed that these snakes ingest prey by rotating the anterior, tooth-bearing halves of the lower jaw rami rapidly in and out of the mouth like a pair of swinging doors (Fig. 1d). This mechanism is made possible by the triple-jointed lower jaw of threadsnakes. In addition to the typical vertebrate jaw joints, which connect the mandible to the rest of the skull and allow the mouth to open like a trapdoor, threadsnakes also have extremely flexible interramal and intramandibular joints (Fig. 1c,d). The interramal joint allows movement between the tips of the mandibular rami, and the intramandibular joints allow the distal halves of the lower jaw rami to rotate backwards into the mouth.

In most other snakes, the interramal and intramandibular joints work passively, allowing the lower jaw to conform more closely to the shape of large vertebrate prey, thereby maximizing the potential gape^{1,2}. In contrast, mandibular raking in *Leptotyphlops* involves active flexion of the intramandibular joints (Fig. 1d). Furthermore, the structure of these joints allows a much greater degree of mandibular flexion than is

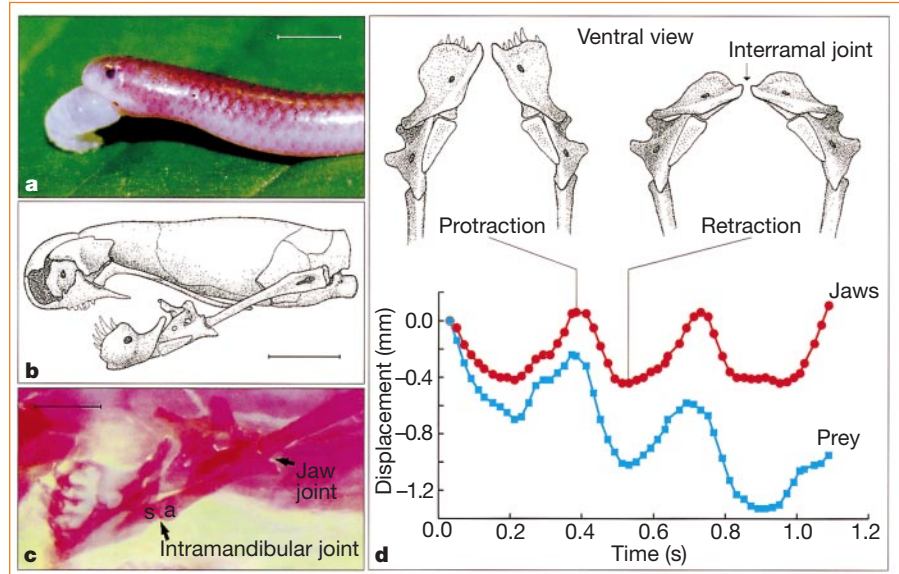


Figure 1 Morphology and function of the feeding apparatus in *Leptotyphlops dulcis*. **a**, An adult swallowing an ant larva. **b**, Left lateral view of the skull. **c**, Left lateral view of the lower jaw of a cleared and alizarin-stained specimen, showing the well developed intra-mandibular joint formed by the contact of the splenial (s) and angular (a) bones. Note the division between the anterior and posterior mandibular segments. **d**, Kinematic plots of jaw (red) and prey (blue) movements for an adult ingesting a large ant pupa, showing the rapid protraction–retraction cycles of the jaws and resultant movement of the prey. Although each jaw protraction results in slight slip-page of the prey, the prey movements associated with jaw retraction are greater, resulting in net transport to the oesophagus. The lower jaw is drawn in ventral aspect to show the mechanics of jaw protraction and retraction. Positions of mandibular elements during retraction were estimated from high-speed video sequences (see Supplementary Information) and drawn from cleared and stained specimens. Scale bars: **a**, 2.5 mm; **b**, 1 mm; **c**, 0.5 mm.

possible in other snakes. In most snakes, the anterior and posterior segments of each mandibular ramus fit closely together in an interdigitating fashion, and flexion at the intramandibular joint is limited by ligaments along the lateral surface of the lower jaw. In *Leptotyphlops*, however, there is a wide gap dorsally between the anterior and posterior mandibular segments, and the ventral articulation between the anterior splenial and posterior angular bones is highly mobile^{6,7} (Fig. 1c). These features make the lower jaw of threadsnakes extremely mobile and allow the transversely oriented mandibular tooth rows (the only teeth in the skull) to be rotated backwards, raking prey into the oesophagus.

Mandibular raking allows threadsnakes to transport prey more rapidly than typical snakes, with cycles of mandibular protraction and retraction often occurring at frequencies exceeding 3 Hz (Fig. 1d). The speed of mandibular raking may be related to the hazardous foraging strategy of threadsnakes. To obtain sufficient quantities of ant brood, the snakes must invade nests that are tenaciously defended by worker ants. Large, aggressive ants can seriously injure or even kill these small snakes⁸, which as adults are usually less

than 2.5 mm in diameter and 1.5 g in weight. Although the anal-gland secretions of *Leptotyphlops* can repel some species of ant⁹, this effect is variable, and selection is likely to favour individuals that can feed quickly, minimizing the time spent exposed to attack.

In contrast, most other snakes engulf their relatively large vertebrate prey by using a 'ptyergoid walk', in which the highly mobile, tooth-bearing elements of the upper jaws are alternately ratcheted over the surface of the prey, thereby 'walking' the snake over and around its prey². It was thought that understanding the feeding mechanisms of scolecophidian snakes (Leptotyphlopidae and two other families of small snakes that have similar feeding habits) may shed light on the origins of the ptyergoid walk, as Scolecophidia occupies a phylogenetic position between 'lizards' and the clade containing all other living snakes (Alethinophidia)^{5,10}. However, the mandibular raking used by *Leptotyphlops* is unlike any feeding mechanism known among lizards or snakes, with prey transport entirely dependent on active flexion of the lower jaw.

Mandibular raking therefore seems to be a uniquely derived feeding mechanism in

Leptotyphlopidae, and is unlikely to represent the primitive feeding mode in snakes, despite the phylogenetically basal position of Scolecophidia.

N. J. Kley, E. L. Brainerd

Organismic and Evolutionary Biology Program and
Department of Biology, University of Massachusetts,
Amherst, Massachusetts 01003-5810, USA
e-mail: kley@bio.umass.edu

1. Gans, C. *Am. Zool.* **1**, 217–227 (1961).
2. Cundall, D. in *Snakes: Ecology and Evolutionary Biology* (eds Seigel, R. A., Collins, J. T. & Nowak, S. S.) 106–140 (Macmillan, New York, 1987).

3. Punzo, F. J. *Herpetol.* **8**, 153–156 (1974).
4. Cundall, D. & Rossman, D. A. *Zool. J. Linn. Soc.* **109**, 235–273 (1993).
5. Greene, H. W. *Snakes: The Evolution of Mystery in Nature* (Univ. California Press, Berkeley, 1997).
6. List, J. C. *Illinois Biol. Monogr.* **36**, 1–112 (1966).
7. McDowell, S. B. Jr & Bogert, C. M. *Bull. Am. Mus. Nat. Hist.* **105**, 1–142 (1954).
8. Webb, J. K. & Shine, R. *Anim. Behav.* **45**, 1117–1126 (1993).
9. Gehlbach, F. R., Watkins, J. F. II & Reno, H. W. *Bioscience* **18**, 784–785 (1968).
10. Cundall, D., Wallach, V. & Rossman, D. A. *Zool. J. Linn. Soc.* **109**, 275–299 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Genetic techniques

A universal marker for transgenic insects

Genetic manipulation of insects and other arthropods may enable better control strategies to be developed against agricultural pests and disease vectors. Transposon-based transformation techniques have been implemented in *Drosophila*¹ and other insects² such as medflies^{3,4} and mosquitoes^{5,6}. A major obstacle in the use of these transposons, however, has been the difficulty in obtaining marker genes that will allow easy and reliable identification of transgenic animals. Here we describe a marker system that is suitable for following gene transfer in most arthropods and in many other phyla.

Species-specific transformation markers can be generated by isolating visible mutations in the species of interest, cloning the corresponding gene, and then rescuing the mutant phenotype by incorporating a wild-type copy of the gene through transformation. But this procedure is laborious, with every species needing the same investment.

A universal marker that could be used to follow gene transfer in any species is the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria*, as it is active across the animal and plant kingdoms⁷. However, GFP requires a strong promoter to enable single-copy insertions to be detected. Activation should be in a spatial pattern so the transgene signal can be distinguished from common autofluorescence.

We find that an artificial promoter containing three binding sites for Pax-6 homodimers in front of a TATA box (3xP3)⁸ is hyperactive, regionally restricted and universal. This promoter can drive expression of an enhanced GFP variant (EGFP)⁷ in the eye of the fruitfly *Drosophila melanogaster* (Fig. 1a) and in the flour beetle *Tribolium castaneum* (Fig. 1b), which are from lineages that separated about 250 million years ago. The evolutionary conservation and the 'master regulatory' function of Pax-6 in the eye development of insects and vertebrates⁹ means that the 3xP3 promoter should be

active in any photoreceptor cell. The small size of the marker gene (1.3 kilobases) allows for small transposon constructs, resulting in high transformation rates.

We constructed three vectors based on the *Hermes*¹⁰, *piggyBac*¹¹ and *mariner*¹² transposons, each carrying the 3xP3-EGFP marker. Together with helper plasmids to provide the respective transposases^{4–6}, these vectors were microinjected into *Drosophila* eggs of a strain mutant for the *white* gene (vector and helper plasmids at 500 and 300 ng μl^{-1} , respectively). We obtained transgenic lines displaying strong fluorescence with transformation efficiencies of 4% for *mariner*, 50% for *Hermes* and 35% for *piggyBac* (the percentage of fertile injection survivors producing fluorescent offspring). In parallel, we microinjected the *Hermes* and *piggyBac* vectors into the posterior pole of *Tribolium* eggs from a strain lacking eye pigmentation (*pearl* mutants). We obtained transgenic beetle lines with frequencies of 1% for *Hermes* and 60% for *piggyBac*.

The transgenes seem to be stably integrated into the genome as they have been inherited over at least six generations. We

detected strong fluorescence for both species, even after outcrossing to wild-type strains (Fig. 1c,d). This result shows that our marker can also be detected in the presence of eye pigments. In both species, all photoreceptor cells express EGFP, including cells in the eyes of larvae (Fig. 1e, f), pupae and adults, and in the ocelli of *Drosophila*.

The transposon-mediated generation of transgenic beetles provides a powerful new technique for studying a large group of pest species. Moreover, owing to its artificial origin⁸, 3xP3-EGFP probably does not require any other host-specific factors and is therefore a universal marker that should function in all animals that have eyes. With a set of promiscuous transposon vectors, such a system can be used to study almost any species, not just established model organisms, in comparative biological and functional evolutionary studies.

Expression in the eyes allows the signal to be seen in animals that do not have transparent cuticle, and transgenic animals can be identified as larvae, pupae and adults. The system can be applied to competitive wild-type strains, rather than just potentially labile mutant lines, making it suitable for pest-management programmes. The use of arthropod transgenics might become so widespread that new regulatory efforts may be needed to oversee the number of species that can be transformed.

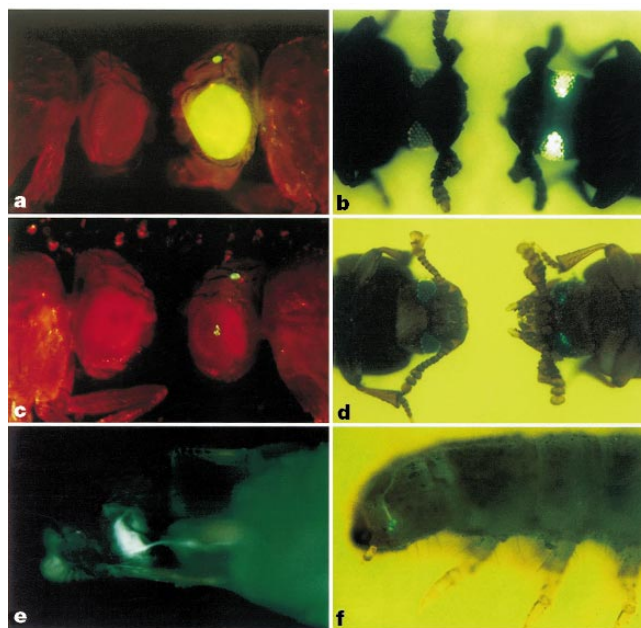
Andreas J. Berghammer*, Martin Klingler*, Ernst A. Wimmer†

*Zoologisches Institut, Universität München,
Luisenstrasse 14, 80333 München, Germany

†Lehrstuhl für Genetik, Universität Bayreuth,
Universitätsstrasse 30, 95447 Bayreuth, Germany
e-mail: ernst.wimmer@uni-bayreuth.de

1. Rubin, G. M. & Spradling, A. C. *Science* **218**, 348–353 (1982).
2. O'Brochta, D. A. & Atkinson, P. W. *Insect Biochem. Mol. Biol.* **26**, 739–753 (1996).

Figure 1 Transgenic fruitflies and flour beetles identified by the universal marker 3xP3-EGFP. **a–d**, Left, non-transgenic controls; right, transgenic individuals marked by 3xP3-EGFP. **a**, In a *Drosophila white* mutant background, the complete compound eye and the ocelli fluoresce. **b**, In a *Tribolium pearl* mutant background, all ommatidia fluoresce. **c**, In a *Drosophila white*+/+ background, fluorescence is seen in the ocelli and the pseudopupil of the compound eye. **d**, In a *Tribolium pearl*+/+ background, fluorescence can be seen only in the ommatidia that point straight towards the observer. **e, f**, In *Drosophila* (**e**) and *Tribolium* (**f**) larvae, EGFP can be seen in the photoreceptors of the larval eyes and in the optic nerve.



3. Loukeris, T. G. *et al. Science* **270**, 2002–2005 (1995).
4. Handler, A. M., McCombs, S. D., Fraser, M. L. & Saul, S. H. *Proc. Natl Acad. Sci. USA* **95**, 7520–7525 (1998).
5. Coates, C. J., Jasinskiene, N., Miyashiro, L. & James, A. A. *Proc. Natl Acad. Sci. USA* **95**, 3748–3751 (1998).
6. Jasinskiene, N. *et al. Proc. Natl Acad. Sci. USA* **95**, 3743–3747 (1998).
7. Tsien, R. Y. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
8. Sheng, G., Thouvenot, E., Schmucker, D., Wilson, D. S. & Desplan, C. *Genes Dev.* **11**, 1122–1131 (1997).
9. Callaerts, P., Halder, G. & Gehring, W. J. *Annu. Rev. Neurosci.* **20**, 483–532 (1997).
10. Warren, W. D., Atkinson, P. W. & O'Brochta, D. A. *Genet. Res.* **64**, 87–97 (1994).
11. Cary, L. C. *et al. Virology* **172**, 156–169 (1989).
12. Medhora, M. M., MacPeck, A. H. & Hartl, D. L. *EMBO J.* **7**, 2185–2189 (1988).

Development

Ubiquitin tag for sperm mitochondria

Like other mammals, humans inherit mitochondria from the mother only, even though the sperm contributes nearly one hundred mitochondria to the fertilized egg. In support of the idea that this strictly maternal inheritance of mitochondrial DNA arises from the selective destruction of sperm mitochondria^{1,2}, we show here that sperm mitochondria inside fertilized cow and monkey eggs are tagged by the recycling marker protein ubiquitin³. This imprint is a death sentence that is written during spermatogenesis and executed after the sperm mitochondria encounter the egg's cytoplasmic destruction machinery.

Figure 1 shows the ubiquitination of sperm mitochondria in fertilized eggs from rhesus monkeys (Fig. 1a; 18 hours post-insemination) and cows (Fig. 1b; 16 hours). At first mitosis, the mitochondrial sheath of the ubiquitinated sperm is evident at one spindle pole (Fig. 1c; 24 hours). The signal typically disappears between the four-cell and the eight-cell stage of development.

The ubiquitination of sperm mitochondria also occurs in the male reproductive tract. The ubiquitinated sites appear to be masked by disulphide bonds during passage through the epididymis. Spermatocyte mitochondria are already ubiquitinated, and most mitochondria are lost to the discarded residual body during spermiation. Mature sperm found on the egg surface and just inside the egg cytoplasm after incorporation are not detectable by ubiquitin antibodies, although several ubiquitinated proteins are found in sperm extracts following the reduction of disulphide bonds. After incorporation and decondensation of sperm within the egg⁴, ubiquitinated sperm mitochondria are again apparent. Sperm mitochondria are lost before, or during, the third embryonic cleavage in mammals^{5–7}.

Remarkably, sperm mitochondria persist in a bovine interspecies cross between the domestic cow and the Asian wild gaur,

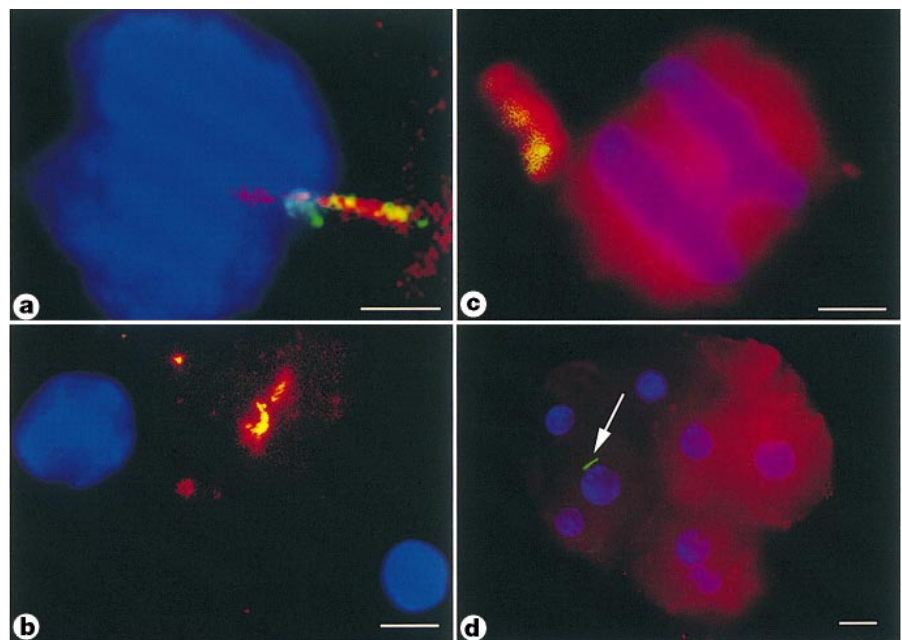


Figure 1 Ubiquitination of sperm mitochondria in mammalian zygotes and embryos. **a–c**, Binding of ubiquitin (red, visualized by anti-ubiquitin antibodies) to the sperm mitochondria (green/yellow, vital dye MitoTracker[®]) within a pronucleate-stage rhesus-monkey zygote (**a**), a bovine pronuclear zygote (**b**) and an intraspecific (domestic cow × domestic bull) embryo during its first mitotic division (**c**). Sperm mitochondria are not found in intraspecific bovine embryos at the eight-cell stage. **d**, In contrast, the intact mitochondrial sheath (green, white arrow) is found in the hybrid eight-cell embryo from domestic cow oocytes fertilized by the sperm of a wild gaur bull. DNA is visualized in blue within the male and female pronuclei (sperm and egg nuclei, respectively, in **a** and **b**), and as the mitotic chromosomes are aligned and separating during first anaphase (**c**) and eight-cell stage embryonic nuclei (**d**). Scale bars: **a–c**, 5 µm; **d**, 20 µm.

as they do in murine hybrids^{8,9}. Sperm mitochondria from wild gaurs (gift from N. Loskutov) persist beyond the third embryonic division (Fig. 1d) after incorporation into cow eggs and are not ubiquitinated, as no antibody binding is detectable. The inheritance of sperm and egg mitochondrial DNA is thought to be solely sex specific, but if it is also species specific, then evolutionary clocks based on mitochondrial DNA may need to be recalibrated.

We propose that the uniparental ubiquitination of mitochondria that occurs during spermatogenesis selectively earmarks sperm mitochondria for degradation by the embryo's proteasomes and lysosomes, for example, during the preimplantation stage. In yeast, ubiquitin mutations interfere with mitochondrial segregation¹⁰, and prohibitin¹¹ (a mitochondrial inner-membrane protein) is one of the likely ubiquitin substrates in mammalian sperm.

Learning more about this destruction pathway is important for understanding inherited mitochondrial diseases, and for evaluating the safety and efficacy of cloning, oocyte cytoplasm 'donation therapy', and fertilization assisted by spermatid or sperm microinjection. The high death rate observed in cloned offspring^{12,13} could be due to heteroplasmy, the condition of mismatched mitochondria.

A report of dissimilar mitochondrial DNA in the first babies born after egg cytoplasmic donation¹⁴ is alarming. Because mitochondrial pathways are disrupted

during cloning and some infertility treatments, foreign mitochondria introduced into the egg may reproduce semiautonomously unless they are properly recognized and processed. The implications of paternal mitochondria being selectively tagged for extinction may force a re-evaluation of procreative strategies.

Peter Sutovsky, Ricardo D. Moreno*, João Ramalho-Santos†, Tanja Dominko, Calvin Simerly, Gerald Schatten

Oregon Regional Primate Research Center and Departments of Cell-Developmental Biology and Obstetrics-Gynecology, Oregon Health Sciences University, 505 NW 185th Avenue, Beaverton, Oregon 97006, USA
e-mail: schatten@ohsu.edu

*International Center for Cancer Research and Developmental Biology, Santiago, Chile

†Center for Neuroscience, Department of Zoology, University of Coimbra, Coimbra, Portugal

1. Hutchinson, C. A., Newbold, J. E. & Potter, S. S. *Nature* **251**, 536–538 (1974).
2. Giles, R. E., Blanc, H., Cann, H. M. & Wallace, D. C. *Proc. Natl Acad. Sci. USA* **77**, 6715–6719 (1980).
3. Ciechanover, A. *Cell* **79**, 13–21 (1994).
4. Perreault, S. D., Wolf, R. A. & Zirkin, B. R. *Dev. Biol.* **101**, 160–167 (1984).
5. Szöllösi, D. J. *Exp. Zool.* **159**, 367–378 (1965).
6. Sutovsky, P., Navara, C. S. & Schatten, G. *Biol. Reprod.* **55**, 1195–1205 (1996).
7. Cummins, J. M., Wakayama, T. & Yanagimachi, R. *Zygote* **5**, 301–308 (1997).
8. Gyllenstein, U., Wharton, D., Josefsson, A. & Wilson, A. C. *Nature* **352**, 255–257 (1991).
9. Kaneda, H. *et al. Proc. Natl Acad. Sci. USA* **92**, 4542–4546 (1995).
10. Fisk, H. A. & Yaffe, M. P. *J. Cell Biol.* **145**, 1199–1208 (1999).

11. Choongkittaworn, N. M., Kim, K. H., Danner, D. B. & Griswold, M. D. *Biol. Reprod.* **49**, 300–310 (1993).
12. Kato, Y. *et al. Science* **282**, 2095–2098 (1998).
13. Renard, J. P. *et al. Lancet* **353**, 1489–1491 (1999).
14. Barrit, J. *et al. Fertil. Steril.* **72**, 31–32 (1999).

Oceanography

Iron, nitrogen and phosphorus in the ocean

It has been proposed that widespread deficits of nitrate in the ocean, like those observed today, are caused by iron limitation of marine nitrogen fixation¹. That is, only when iron is sufficiently abundant to satiate nitrogen fixers will the ratio of nitrate to phosphate in the ocean increase to 16, the average for phytoplankton. Tyrrell² developed a simple two-box model of oceanic nitrogen and phosphorus cycles to describe the regulation of both nitrate and phosphate concentrations in the global ocean. His criterion for nitrate deficit in the ocean, a molar ratio of N:P in surface waters (R_S) of less than 16, is satisfied without recourse to iron limitation, calling into question Falkowski's proposal¹ about the biogeochemical significance of iron limitation as it relates to nitrogen fixation and oceanic levels of nitrogen and phosphorus. Here I show that small changes in the assumptions of Tyrrell's model, well within acknowledged uncertainty, can lead to values of R_S greater than 16. Consequently, the consistency of the model with the observed distributions of nutrients in the ocean is uncertain, and the influence of iron may still be considered important.

In Tyrrell's model², competition between nitrogen-fixing and other phytoplankton controls the level of nitrate, continually pushing molar concentrations to slightly less than 16 times those of phosphate. Results of the model show that phosphate is the ultimate limiting nutrient because extra phosphate in the system supports the proliferation of nitrogen fixers that can add new nitrogen to the ocean. Even when subjected to extensive sensitivity analysis², the model consistently predicts a deficit of nitrate in the surface layer, defined by Tyrrell as $R_S < 16$.

Further calculations (Fig. 1) indicate that the model's prediction of a nitrate deficit in surface waters of the ocean is uncertain. R_S is very sensitive to chosen values for the Michaelis–Menten half-saturation constants for the growth of phytoplankton on nitrate ($K_S[\text{NO}_3]$) and phosphate ($K_S[\text{PO}_4]$) (N_H and P_H , respectively, in Tyrrell's notation). A 20% increase of $K_S[\text{NO}_3]$ to 0.6 mmol m⁻³ from the assumed 0.5 mmol m⁻³ obliterates the predicted nitrate deficit, bringing R_S to 16. The reported² uncertainty in $K_S[\text{NO}_3]$ is 0.1 to

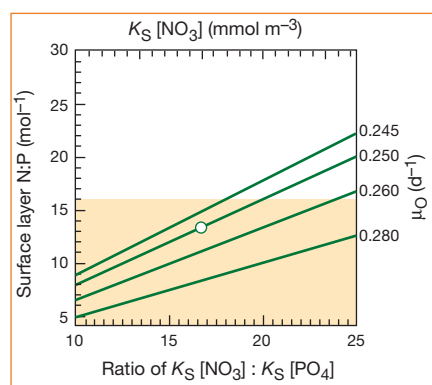


Figure 1 Influence of assumed Michaelis–Menten half-saturation constants and phytoplankton growth rates on steady-state solutions for R_S , the molar ratio of nitrate to phosphate in the surface layer of the ocean (model of ref. 2). Nitrate deficit at the surface is indicated by $R_S < 16$ (shaded). The half-saturation constant for growth versus nitrate, $K_S[\text{NO}_3]$, was varied from the specified 0.5 mmol m⁻³ to obtain a range of $K_S[\text{NO}_3]:K_S[\text{PO}_4]$, keeping $K_S[\text{PO}_4]$ constant at the specified 0.03 mmol m⁻³. The solution for R_S against $K_S[\text{NO}_3]:K_S[\text{PO}_4]$ is independent of $K_S[\text{PO}_4]$ (equation 13 in ref. 2). The maximum growth rate of nitrogen-fixing phytoplankton, μ'_{NF} , was maintained at 0.24 d⁻¹, as specified in the model. The maximum growth rate of other phytoplankton, μ'_O , was varied from the original 0.25 d⁻¹, as indicated. The original model solution is identified with the open circle.

4.2 mmol m⁻³ (ref. 3), corresponding to an R_S value between 2.7 and 112 mol mol⁻¹.

Small changes in the maximum growth rate for other phytoplankton, μ'_O (d⁻¹), compared with 0.24 d⁻¹ for nitrogen fixers, also strongly influence R_S (Fig. 1). There is little experimental basis for excluding assumed growth rates that lead to an R_S value of 16.

This simple analysis, based directly on Tyrrell's model, suggests that regulation of oceanic nitrogen fixation by iron cannot be excluded as a potentially important influence on cycles of nutrients and primary productivity in the ocean.

John J. Cullen

Center for Environmental Observation Technology and Research, Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada
e-mail: john.cullen@dal.ca

Tyrrell replies — Cullen shows that changing the values of P_H and N_H ($K_S[\text{PO}_4]$ and $K_S[\text{NO}_3]$ in his notation) in my model² can give rise to a steady-state $[\text{NO}_3]:[\text{PO}_4]$ ratio in surface waters that is greater than or equal to 16. Because of this, and because P_H and N_H are not well known, he questions some of the implications of my model.

Although his analysis is correct, my model must converge to a $[\text{NO}_3]:[\text{PO}_4]$ ratio that is slightly less than the ratio at which NO_3 and PO_4 are equally limiting to growth, regardless of whether the latter ratio is 16 or not. There must therefore be convergence to proximate nitrate (reactive

nitrogen) limitation of surface waters.

Equations (1) and (2) of my model² can be rewritten as

$$d(\text{NF})/dt = (\mu'_{\text{NF}} \text{LP} - M)\text{NF}$$

and

$$d(\text{O})/dt = (\mu'_O \min\{\text{LN}, \text{LP}\} - M)\text{O}$$

where NF and O are the populations of nitrogen-fixing and other phytoplankton, t is time, M is mortality, and LN and LP represent the growth limitations (range, 0 to 1) caused by $[\text{NO}_3]$ and $[\text{PO}_4]$ shortages, respectively, which were expressed as Michaelis–Menten functions in ref. 2 but have been left unspecified here. For steady state, $d(\text{NF})/dt$ and $d(\text{O})/dt$ must both equal zero (both populations are stable in size), and therefore

$$\mu'_{\text{NF}} \text{LP} = M = \mu'_O \min\{\text{LN}, \text{LP}\}$$

or

$$\mu'_{\text{NF}}/\mu'_O = \min\{\text{LN}, \text{LP}\}/\text{LP}$$

Because μ'_{NF} is less than μ'_O (ref. 1), this equation cannot be satisfied (equilibrium cannot occur) unless LN is less than LP; that is, the proximate limiting nutrient is reactive nitrogen.

This proof makes no assumptions about the values of N_H and P_H . As Cullen rightly argues, raising N_H allows convergence to a surface N:P > 16, but then the surface ocean is still most strongly limited by reactive nitrogen, precisely because N_H has been raised. But if the simplifying assumption is made that the Michaelis–Menten half-saturation constant for a nutrient is more or less proportional to the rate at which it needs to be taken up to fuel new growth (in which case, N_H/P_H is about 16), then $\text{LN} < \text{LP}$ also implies that $[\text{NO}_3]:[\text{PO}_4] < 16$ in the surface ocean steady state. Cullen's analysis illuminates the point that a surface ocean could still have nitrogen as the proximate limiting nutrient even if the surface N:P ratio were 100, for instance. It all depends on the values of P_H and N_H , which need to be better constrained.

This analysis confirms my original point: observations from nutrient-enrichment experiments that reactive nitrogen is more limiting to growth than phosphorus in the surface ocean can be reconciled with phosphorus limitation without recourse to the effects of trace metals. If shortages of iron, for instance, further depress nitrogen fixation and reactive nitrogen concentrations in the open ocean, then my model predicts that adding extra iron could cause only a reduction, not full removal, of the proximate nitrogen limitation.

Toby Tyrrell

School of Ocean and Earth Sciences, Southampton Oceanography Centre, University of Southampton, Southampton SO14 3ZH, UK

1. Falkowski, P. G. *Nature* **387**, 272–275 (1997).
2. Tyrrell, T. *Nature* **400**, 525–531 (1999).
3. Goldman, J. C. & Glibert, P. M. in *Nitrogen in the Marine Environment* (eds Capone, D. G. & Carpenter, E. J.) 233–274 (Academic, New York, 1983).

Active site-directed protein regulation

Bostjan Kobe & Bruce E. Kemp

St. Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia

Regulation of protein function is vital for the control of cellular processes. Proteins are often regulated by allosteric mechanisms, in which effectors bind to regulatory sites distinct from the active sites and alter protein function. Intrasteric regulation, directed at the active site and thus the counterpart of allosteric control, is now emerging as an important regulatory mechanism.

Protein function in the cell can be modulated in a number of ways, including noncovalent interactions with regulatory factors, covalent modification (for example, phosphorylation, lipidation) or proteolytic cleavage. A widespread mechanism of the direct control of protein function is 'allosteric' regulation, in which effectors bind to regulatory sites distinct from the active site¹, usually inducing conformational changes that profoundly influence the activity. Allosteric effectors generally bear no structural resemblance to their target protein's substrate. This form of regulation explains how the end products of metabolic pathways can act at early steps of the pathway to exert feedback control. Allosteric regulation has been studied intensively for over three decades and is now considered a classical control mechanism of protein function.

The term intrasteric regulation was introduced to describe autoregulation of protein kinases and phosphatases by internal sequences resembling substrates ('pseudosubstrates') and acting directly at the active site² (Fig. 1). Several lines of indirect evidence supporting intrasteric regulation were initially obtained using various biochemical approaches. For example, synthetic peptides corresponding to the putative pseudosubstrate regulatory sequences acted as inhibitors; and removal of the regulatory sequences by truncation mutagenesis or limited proteolysis reversed the autoinhibition, as did antibodies directed towards the regulatory sequence³. Unequivocal evidence of intrasteric control, however, has only become available through structural studies of several autoinhibited enzymes; these include the crystal structures of protein kinase A (PKA) with the bound peptide inhibitor⁴, twitchin kinase⁵, titin kinase⁶, calmodulin-dependent protein kinase-1 (CaMK-1)⁷ and the protein phosphatase calcineurin⁸.

But is the scope of intrasteric control restricted merely to protein kinases and phosphatases? It now emerges that this form of control extends to diverse enzyme classes and beyond, to receptors and protein-targeting domains. Intrasteric control can be used for both autoregulation, either through intramolecular (Fig. 1a) or intermolecular interactions (such as *trans* interactions between monomers in an oligomer; Fig. 1b), and more general regulation through intermolecular interactions (Fig. 1c). Here we focus on structurally characterized examples in which control is mediated by active site-directed, intrasteric autoregulatory sequences (IARS) (Table 1).

Intrasteric interactions typically inhibit protein function. But how is the autoinhibition reversed? Mechanisms used for activation include protein activators or ligands, phosphorylation, proteolysis, reduction of disulphide bonds and combinations of these (Table 1). Intrasteric inhibition requires allosteric control to switch protein function back on (Fig. 1). This combination of intrasteric and allosteric controls provides a powerful and flexible mechanism to control some of the most complex cellular processes.

IARS-dependent intrasteric interactions

Protein kinases. Protein phosphorylation is the most common mechanism of cellular regulation, and, to ensure signalling fidelity, protein kinases (and phosphatases) must be tightly regulated. Many exist in latent forms requiring activation by specific modulators. A large subfamily of kinases are activated by calcium-binding proteins

(such as myosin-light chain kinase), or directly by calcium (such as plant Ca-dependent protein kinase family; these kinases contain a Ca-binding domain contiguous with the kinase domain). Crystal structures of fragments of twitchin kinase, a giant myosin-associated kinase, show that a carboxy-terminal IARS threads through the active-site cleft between the two lobes of the protein kinase domain, making many contacts with the peptide substrate binding site and other regions essential for catalysis and ATP binding^{5,9} (Fig. 2a). The binding site of the activator S100A1 has been mapped to one portion of the autoinhibitory sequence, and reverses these interactions to permit substrate access^{9,10}. A similar mechanism of inhibition occurs in the related giant titin kinase; however, here a combination of phosphorylation and calmodulin (CaM) binding (in a site analogous to S100A1-binding site in twitchin) activates the enzyme⁶. The more distantly related CaMK-1 shows a modified mode of inhibition in which the IARS exits the active site in the opposite direction to that in twitchin kinase⁷. Although there is no three-dimensional information for other calcium-dependent protein kinases, these are expected to exhibit regulatory mechanisms that are related to those seen in twitchin, titin and CaMK-1. Other protein kinase families predicted to be intrastERICALLY regulated include the protein kinase C family and cGMP-dependent protein kinase, in which the IARS are located amino-terminal to the catalytic domain and the activators are phospholipids and diacylglycerol, and cGMP, respectively³.

The crystal structure of the catalytic domain of the insulin receptor tyrosine kinase revealed a more subtle autoinhibitory mechanism with a tyrosine residue bound to the active site¹¹. This

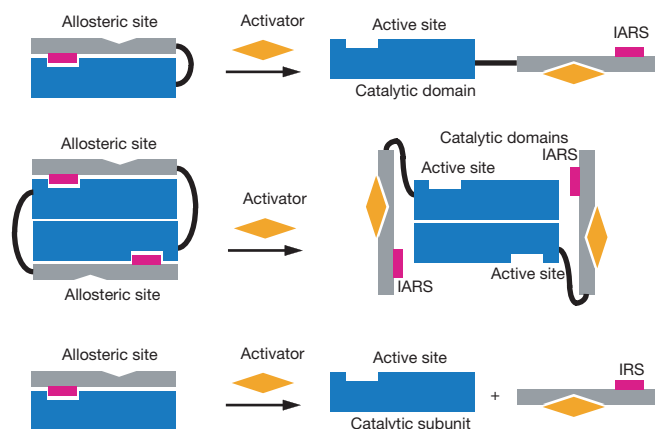


Figure 1 Representation of intrasteric regulation. **a**, Intrasteric regulation of enzymes by intramolecular interactions, through an intrasteric autoregulatory sequence (IARS) that is contiguous with the catalytic domain. The enzyme is maintained in an inactive state through the binding of the IARS in an intramolecular fashion that masks the active site. Allosteric activation by an activatory ligand or protein results in the release of the IARS from the active site. **b**, Intrasteric regulation of homomeric enzymes by intermolecular interactions, through an IARS in *trans*. The IARS interacting with one subunit is a part of another subunit. **c**, Intrasteric regulation of heteromeric enzymes by intermolecular interactions, through an intrasteric regulatory sequence (IRS) on a distinct subunit.

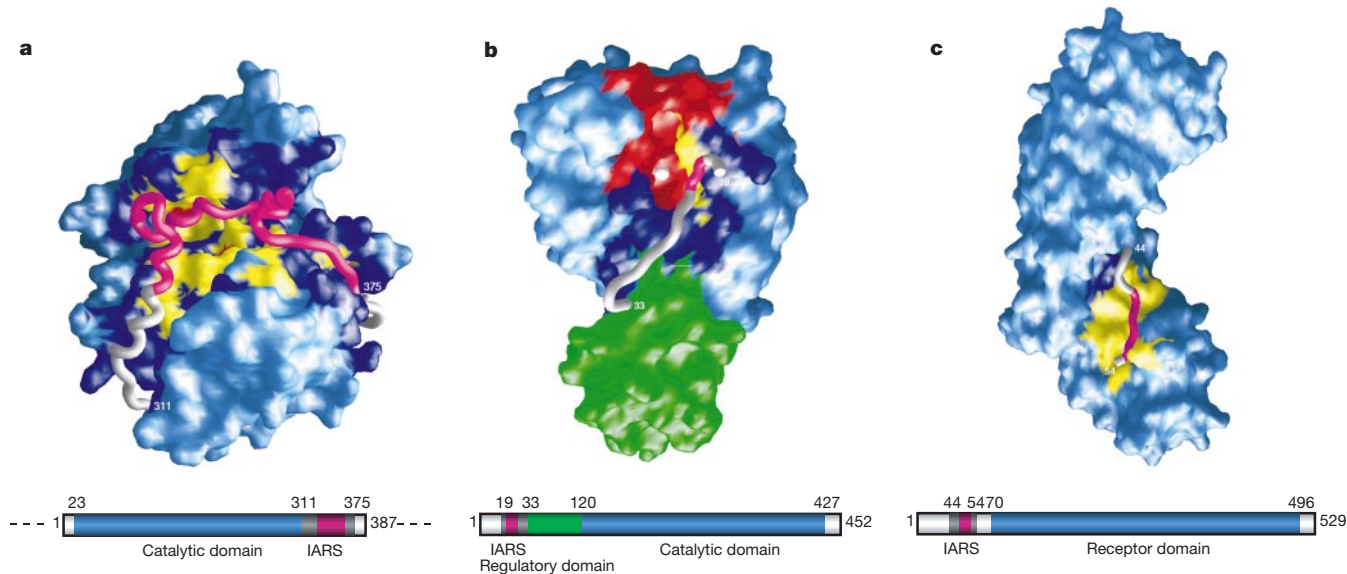


Figure 2 Three-dimensional (3D) structures of representative proteins under intrasteric control through active site-directed autoregulatory sequences. Although IARS can be coarsely identified through limited proteolysis or truncation and site-directed mutagenesis, only 3D structure determination can delineate the contacts with the active site in detail. We define the IARS as comprising both the sequences interacting with the active site (magenta) and the linker sequence (grey) tethering this segment to other globular domains. **a**, 3D structure of the twitchin kinase fragment from *Aplysia* (PDB accession code 1KOB). The IARS is shown in a tube representation, and the rest of the protein in a surface representation (the active-site regions that interact with the IARS are yellow; the active-site regions that do not interact with the IARS are red; the regions not in

the active site but interacting with the IARS are dark blue; and the rest of the surface is light blue). Generated with GRASP⁴⁶. Bottom, diagram of protein sequence. **b**, 3D structure of a monomer of phenylalanine hydroxylase (PDB accession code 1PHZ), coloured and drawn as in **a**. The regulatory domain is shown in green, and the active-site iron atom is shown as a white sphere. All residues lining the active-site cavity have been considered as active-site residues. Only a minor portion of the active site interacts with the IARS. **c**, 3D structure of importin-α (PDB accession code 1IAL), coloured and drawn as in **a**. The IARS (magenta) acts as a pseudo-NLS contacting the major NLS-binding site (yellow).

tyrosine may be considered as a transient pseudosubstrate and is ultimately autophosphorylated in response to insulin binding to the extracellular part of the receptor. Phosphorylation of this and two other tyrosine residues results in a rearrangement, allowing access to the active site¹². The active site is similarly blocked by a tyrosine in the inactive structure of the serine/threonine kinase mitogen-activated protein (MAP) kinase extracellular signal-regulated kinase (ERK)2¹³. The tyrosine is one of two residues phosphorylated by a distinct upstream kinase to yield the active enzyme¹⁴. **Protein phosphatases.** Autoinhibition of the serine/threonine phosphatase calcineurin (CaN) is caused by an 18-residue sequence C-terminal to the catalytic domain that lies over the substrate-binding cleft of the A chain (CaNA)⁸. It is activated by calcium

binding to the CaM-like B chain (CaNB), and by calcium/CaM binding to CaNA. CaNB and CaM bind to distinct regions in the sequence between the catalytic domain and the autoinhibitory sequence of CaNA. Although three-dimensional information is still lacking, other serine/threonine protein phosphatases are predicted to be intrasterically regulated by both noncatalytic segments and various binding proteins (see above); in some cases, pseudosubstrate sequences have been recognized^{2,15}. The crystal structures of tyrosine phosphatase SHP-2 (ref. 16) and the membrane proximal domain of receptor protein tyrosine phosphatase-α (RPTP-α)¹⁷ revealed two further intrasteric mechanisms of inhibition. In SHP-2, the phosphatase active site is blocked by a noncatalytic targeting, Src homology (SH)2 domain

Table 1 Examples of regulation mediated by intrasteric autoregulatory sequences

Protein class	Protein	IARS location	Pseudosubstrate	Activation
Protein kinases	Twitchin kinase	C-terminal	Partial	S100A1 binding
	Titin kinase	C-terminal	Partial	CaM binding + phosphorylation
	CaMK-1	C-terminal	Partial	CaM binding + phosphorylation
	Insulin receptor kinase	Internal	Substrate	Phosphorylation
	MAP kinase ERK2	Internal	Not apparent	Phosphorylation
Protein phosphatases	Calcineurin	C-terminal	Partial	CaM binding
	Receptor PTP-α	Other subunit in homodimer	Not apparent	Phosphorylation
	SHP-2	N-terminal	Not apparent	Phosphoprotein binding
Proteinases	Serine proteinases	N-terminal	Not apparent	Proteolysis
	Aspartic proteinases	N-terminal	Not apparent	Proteolysis
	Cysteine proteinases	N-terminal	Not apparent	Proteolysis
	Zinc metalloproteinases	N-terminal	Not apparent	Proteolysis
	Multi-molecular assemblies	N-terminal	Not apparent	Proteolysis
Metabolic enzymes	Phenylalanine hydroxylase	N-terminal	Not apparent	Phenylalanine binding
	Chloroplast malate dehydrogenase	C-terminal	Not apparent	Oxidation
	Yeast glycogen phosphorylase	N-terminal	Not apparent	Phosphorylation
Transport receptors	Importin-α	N-terminal	Yes	Importin-β binding
Targeting domains	Src, Hck SH2	C-terminal	Yes	Dephosphorylation
	Src, Hck SH3	C-terminal	Yes	SH3 ligands

(the phosphopeptide-binding cleft of the SH2 domain is simultaneously disrupted by this interaction), and activation occurs through phosphopeptide binding by the SH2 domain. In RPTP- α , one molecule in the homodimer sterically blocks the active site in the other, forcing it to adopt an open, inactive conformation, and the dissociation is proposed to occur upon phosphorylation. This regulatory mechanism may be physiologically significant¹⁸. There is strong evidence that receptor dimerization is important in signal transduction events¹⁹ and intrasteric regulation may be exploited in a number of these examples.

Proteinases. Proteinases have diverse biological functions from catabolic digestion to blood coagulation, apoptosis, cleavage of viral precursor proteins to functional units, and pattern formation in multicellular organisms. They are synthesized as inactive precursors ('zymogens') to prevent unwanted protein degradation and to enable regulation of proteolytic activity with respect to space and time²⁰. Proteolytic activity is further controlled by specialized proteins that specifically inhibit the active enzymes.

The structural basis of zymogen autoinhibition was first shown in crystallographic studies of chymotrypsinogen²¹ and trypsinogen^{22,23}, and has since been shown in all classes of proteolytic enzyme²⁰. The inhibitory 'activation segments' in zymogens are usually N-terminal extensions of the mature enzymes that sterically block the active site. They can also be insertions within the sequence of the mature enzyme, but, significantly, they never occur at the C termini, which presumably prevents proteolytic activity during polypeptide synthesis. These segments often have other roles such as folding, stability and/or intracellular sorting. Because the most common mechanism for conversion to active enzymes is limited proteolysis of the inhibitory segment, they are also termed 'prosegments'. The inhibitory segments tend not to mimic the peptide substrates closely, as they have evolved to prevent self-cleavage before the appropriate signal. Consequently, the prosegments of some enzymes (such as papain-like cysteine proteases and pro-stromelysin-1) extend across the active site in the reverse direction to that required for cleavage, whereas the prosegments of others interact with the active site in conformations (such as through a helix in aspartic proteinases) distinct from the extended structures of substrate peptides.

Zymogens have various activation mechanisms which may involve accessory molecules, or simply a drop in pH to initiate autocatalysis. The prosegments are usually degraded after activation, presumably to ensure that the conversion is irreversible and that the prosegments do not continue to act as inhibitors. This irreversible activation by proteolysis distinguishes zymogens from other intrasterically regulated proteins, in which the activation is reversible.

Mechanisms of inhibition of mature enzymes by protein proteinase inhibitors (such as the bovine pancreatic trypsin inhibitor (BPTI) family, serpins or cystatins) are often distinct from the strategies used by prosegments in zymogens^{20,24}. Some very closely mimic substrates and function by trapping the enzyme into a slow and inefficient cleavage reaction.

Metabolic enzymes. Phenylalanine hydroxylase (PAH) is tightly regulated by the substrates phenylalanine and tetrahydrobiopterin, and also by phosphorylation. The crystal structure of the inactive form of PAH shows an N-terminal IARS containing the phosphorylation site which extends from the core of the regulatory domain over the active-site pocket (Fig. 2b)²⁵. Activation may occur by phenylalanine binding to an allosteric site, inducing conformational changes that lead to the removal of the autoinhibitory region from the active site. Phosphorylation in the IARS facilitates this process, but cannot by itself activate the enzyme.

Intrasteric control is also responsible for the light regulation of chloroplast nicotinamide adenine dinucleotide phosphate (NADP)-dependent malate dehydrogenase. Light regulation is mediated by thioredoxin-catalysed reduction and re-oxidation of cysteine residues. The structure of the inactive, oxidized protein

shows a C-terminal IARS that is constrained by a disulphide bond and forced to fold into the active site^{26,27}.

In yeast phosphorylase, an N-terminal IARS blocks the active site of the adjacent monomer in the dephosphorylated form; and phosphorylation of the IARS causes it to shift from the active site²⁸. In contrast, mammalian phosphorylase is regulated by a mechanism not involving intrasteric interactions by an IARS²⁹.

Transport receptors. Importin- α is the nuclear import factor that recognizes classical monopartite and bipartite nuclear localization signals (NLSs). Mouse importin- α has an N-terminal IARS containing a pseudo-NLS bound to the NLS-binding-site on the large C-terminal receptor domain (Fig. 2c)³⁰. This structure represents the nuclear, low-affinity form of the nuclear import receptor for NLS binding. Conversion to the high-affinity form found in the cytoplasm occurs through binding of the activator protein, importin- β , to a sequence overlapping the pseudo-NLS; importin- β binding markedly alters the conformation of the IARS and presumably exposes the NLS-binding site³¹. Importin- α is a classical example in which biochemical approaches have led to conflicting conclusions regarding intrasteric control. Similar approaches using proteolytic digestion, truncation mutagenesis and synthetic peptides suggested the presence of an autoinhibitory sequence to some authors^{32,33}, but the opposite to others^{34,35}. This underscores, as do studies on protein kinases, the need for three-dimensional information in understanding autoregulation of protein function by intrasteric regulation.

Targeting domains. Many signalling proteins contain protein interaction domains, such as SH2 and SH3, that target these proteins to specific substrates. The crystal structures of the inactive forms of c-Src and the related Hck^{36–39} show that the SH2 and SH3 domains are responsible for kinase inactivation through an allosteric mechanism. From another perspective, however, the binding activity of these domains for other binding partners is autoinhibited by kinase sequences through an intrasteric mechanism. The linker connecting the SH2 and the catalytic domain forms a polyproline II helix and binds to the SH3 domain, whereas the SH2 domain binds to a phosphotyrosine in the C-terminal tail of the protein; both interactions are therefore pseudosubstrate-like with respect to these protein interaction domains. Thus, in a highly reciprocal way, the SH2 and SH3 domains regulate the kinase activity while their own targeting activity is simultaneously modulated by the kinase. Tail dephosphorylation and high-affinity ligands for SH2 and SH3 domains would result in kinase activation. A protein kinase domain has been similarly used to autoinhibit the guanylate cyclase activity of the atrial natriuretic peptide receptor, but further structural studies are required to establish the mechanism⁴⁰.

IARS-independent intrasteric interactions

Alternative intrasteric interactions not mediated by IARSs include active-site modifications, such as intrasteric phosphorylation (inhibition by phosphorylation in the active site), and the binding of inhibitory cellular ligands (both proteinaceous and nonproteinaceous) to the active site. In the cyclin-dependent protein kinases, phosphorylation occurs on two residues in the nucleotide-binding loop by the Wee1 kinase⁴¹, and in *Escherichia coli* isocitrate dehydrogenase (IDH) phosphorylation occurs directly on an active-site residue by the bifunctional IDH kinase/phosphatase⁴². Binding of inhibitors to the active sites and forming intrasteric interactions through intrasteric regulatory sequences is very common in proteinases (BPTI, serpins, cystatins), ribonucleases (mammalian ribonuclease inhibitor, barstar) and protein kinases (protein kinase inhibitor, PKA regulatory subunits).

Conclusions

Our examples of intrasteric regulation confirm predictions that pseudosubstrates act at the active site, and expectations that the contacts made by the inhibitory segments are more complex and do not merely mimic the substrates (compare the yellow and red

regions in Fig. 2). In cases such as zymogens, it is advantageous for the inhibitory segment to be more 'pseudo' than 'substrate', to prevent unwarranted activation.

Although unequivocal evidence of intrasteric control can only be provided through structural studies, other evidence, obtained, for example, by truncation mutagenesis, limited proteolysis, synthetic peptides corresponding to the regulatory sequences and antibodies directed towards the regulatory sequences, can shed further light on the regulatory process. Structural information alone is not sufficient evidence for a biologically relevant regulatory mechanism; intrasteric interactions through C-terminal sequences observed in fragments of chaperones DnaK and Hsc70 may be caused by the truncated protein constructs used^{43,44}, and similar interactions of a C-terminal peptide with an adjacent molecule in the crystals of protein farnesyltransferase⁴⁵ may be driven by the crystallization process rather than by native regulation.

In many cases, the interactions of the inhibitory segment with the rest of the protein do not appear particularly strong. Many proteinase prosegments or portions thereof have high-temperature factors suggesting high mobility, and in importin- α regions flanking the autoinhibitory sequence are disordered. It seems reasonable that this is important in facilitating the activation process, and the nature of the interactions formed by the IARS must clearly reflect the mechanism of activation. As an example, a prominent feature of the pH-activated zymogen prosegment interactions is the critical role of salt bridges. Because the thermodynamics and kinetics of inhibition and activation processes in intrasterically regulated proteins are generally poorly understood, studies of these parameters accompanied by further structural studies of both the inhibited and activated states should represent some of the most important future research directions.

The examples show that intrasteric regulation through IARS can occur in monomeric systems (in which the IARS and the catalytic regions are located within the same polypeptide chain; Fig. 1a) and homomeric systems (in which the IARS is located either in the same monomer as the catalytic domain (*cis*, intramolecular interaction) or within a different monomer of the oligomer (*trans*, intermolecular interaction; Fig. 1b)) (Table 1). Intrasteric interactions can also be achieved in other ways, such as by regulatory sequences contained within transient protein subunits (such regulation can be thought of as analogous to the regulation of homomeric proteins in *trans*, but instead occurring in heteromeric systems; Fig. 1c); covalent modifications of the active-site residues (for example, intrasteric phosphorylation); and binding of cellular effectors to the active site. Intrasteric regulation may be a widespread control mechanism that is prominent in the regulation of cellular functions; it will undoubtedly be found in many more proteins as the repertoire of protein three-dimensional structures expands. Intrasteric control is typically accompanied by allosteric control to turn protein function on; it is this combination of intrasteric and allosteric controls that provides the flexibility to achieve almost any level of regulation of protein function that may be required. □

1. Monod, J., Changeux, J. P. & Jacob, F. Allosteric proteins and cellular control systems. *J. Mol. Biol.* **6**, 306–329 (1963).
2. Kemp, B. E. & Pearson, R. B. Intrasteric regulation of protein kinases and phosphatases. *Biochim. Biophys. Acta* **1094**, 67–76 (1991).
3. Kemp, B. E. *et al.* in *Protein Kinases* (ed. Woodgett, J. R.) 30–67 (IRL, Oxford, 1994).
4. Knighton, D. R. *et al.* Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 414–420 (1991).
5. Hu, S.-H. *et al.* Insights into autoregulation from the crystal structure of twitchin kinase. *Nature* **369**, 581–584 (1994).
6. Mayans, O. *et al.* Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395**, 863–869 (1998).
7. Goldberg, J., Nairn, A. C. & Kuriyan, J. Structural basis for the auto-inhibition of calcium/calmodulin-dependent protein kinase I. *Cell* **84**, 875–887 (1996).
8. Kissinger, C. R. *et al.* Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Nature* **378**, 641–644 (1995).
9. Kobe, B. *et al.* Giant protein kinases: domain interactions and structural basis of autoregulation. *EMBO J.* **15**, 6810–6821 (1996).

10. Heierhorst, J. *et al.* Ca²⁺/S100 regulation of giant protein kinases. *Nature* **380**, 636–639 (1996).
11. Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754 (1994).
12. Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **16**, 5572–5581 (1997).
13. Zhang, F., Strand, A., Robbins, D., Cobb, M. H. & Goldsmith, E. J. Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* **367**, 704–711 (1994).
14. Canagarajah, B. J., Khokhlatchev, A., Cobb, M. H. & Goldsmith, E. J. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* **90**, 859–869 (1997).
15. Barford, D., Das, A. K. & Egloff, M. P. The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 133–164 (1998).
16. Hof, P., Pluskey, S., Dhe-Paganon, S., Eck, M. J. & Shoelson, S. E. Crystal structure of the tyrosine phosphatase SHP-2. *Cell* **92**, 441–450 (1998).
17. Bilwes, A. M., den Hertog, J., Hunter, T. & Noel, J. P. Structural basis for inhibition of receptor protein-tyrosine phosphatase- α by dimerization. *Nature* **382**, 555–559 (1996).
18. Jiang, G. *et al.* Dimerization inhibits the activity of receptor-like protein-tyrosine phosphatase- α . *Nature* **401**, 606–610 (1999).
19. Weiss, A. & Schlessinger, J. Switching signals on or off by receptor dimerization. *Cell* **94**, 277–280 (1998).
20. Khan, A. R. & James, M. N. G. Molecular mechanisms for the conversion of zymogens to activate proteolytic enzymes. *Protein Sci.* **7**, 815–836 (1998).
21. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. & Xuong, N. H. Chymotrypsinogen: 2.5-Ångström crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. *Biochemistry* **9**, 1997–2009 (1970).
22. Kossiakoff, A. A., Chambers, J. L., Kay, L. M. & Stroud, R. M. Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry* **16**, 654–664 (1977).
23. Fehllhammer, H., Bode, W. & Huber, R. Crystal structure of bovine trypsinogen at 1–8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J. Mol. Biol.* **111**, 415–438 (1977).
24. Bode, W. & Huber, R. Natural protein proteinase inhibitors and their interactions with proteinases. *Eur. J. Biochem.* **204**, 433–452 (1992).
25. Kobe, B. *et al.* Structural basis of autoregulation of phenylalanine hydroxylase. *Nature Struct. Biol.* **6**, 442–448 (1999).
26. Carr, P. D., Verger, D., Ashton, A. R. & Ollis, D. L. Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. *Structure* **7**, 461–475 (1999).
27. Johansson, K. *et al.* Structural basis for light activation of a chloroplast enzyme: the structure of sorghum NADP-malate dehydrogenase in its oxidized form. *Biochemistry* **38**, 4319–4326 (1999).
28. Lin, K., Hwang, P. & Fletterick, R. J. Distinct phosphorylation signals converge at the catalytic center in glycogen phosphorylase. *Structure* **5**, 1511–1523 (1997).
29. Johnson, L. N., Barford, D., Owen, D. J., Noble, M. E. & Garman, E. F. From phosphorylase to phosphorylase kinase. *Adv. Second Messenger Phosphoprotein Res.* **31**, 11–28 (1997).
30. Kobe, B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nature Struct. Biol.* **6**, 388–397 (1999).
31. Cignolini, G., Petosa, C., Weis, K. & Müller, C. W. Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399**, 221–229 (1999).
32. Moroiu, J., Blobel, G. & Radu, A. The binding site of karyopherin alpha for karyopherin beta overlaps with a nuclear localization sequence. *Proc. Nat. Acad. Sci. USA* **93**, 6572–6576 (1996).
33. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94**, 193–204 (1998).
34. Gorlich, D., Henklein, P., Laskey, R. A. & Hartmann, E. A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus. *EMBO J.* **15**, 1810–1817 (1996).
35. Weis, K., Ryder, U. & Lamond, A. I. The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import. *EMBO J.* **15**, 1818–1825 (1996).
36. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595–602 (1997).
37. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602–653 (1997).
38. Williams, J. C. *et al.* The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J. Mol. Biol.* **274**, 757–775 (1997).
39. Moarefi, I. *et al.* Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* **385**, 650–653 (1997).
40. Chinkers, M. & Garbers, D. L. The protein kinase domain of the ANP receptor is required for signalling. *Science* **245**, 1392–1394 (1989).
41. Morgan, D. O. Principles of CDK regulation. *Nature* **374**, 131–134 (1995).
42. Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E. Jr & Stroud, R. M. Regulation of an enzyme by phosphorylation at the active site. *Science* **249**, 1012–1016 (1991).
43. Wang, H. *et al.* NMR solution structure of the 21 kDa chaperone protein DnaK substrate binding domain: a preview of chaperone-protein interaction. *Biochemistry* **37**, 7929–7940 (1998).
44. Morshauer, R. C. *et al.* High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. *J. Mol. Biol.* **289**, 1387–1403 (1999).
45. Park, H.-W., Boduluri, S. R., Moomaw, J. F., Casey, P. J. & Beese, L. S. Crystal structure of protein farnesyltransferase at 2.25 Ångström resolution. *Science* **275**, 1800–1804 (1997).
46. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements

We thank J. Heierhorst, I. Jennings and other members of the protein research groups at St. Vincent's Institute for discussions; T. Teh for comments on the manuscript; and P. Carr for unpublished data. The work was supported by the National Health and Medical Research Council, Wellcome Trust, Australian Research Council and National Heart Foundation; B.K. is a Wellcome Senior Research Fellow in Medical Science in Australia; B.E.K. is an NHMRC Fellow.

Correspondence and requests for materials should be addressed to B.K. (e-mail: b.kobe@medicine.unimelb.edu.au).

Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution

C. Roy D. Lancaster*, Achim Kröger†, Manfred Auer*‡ & Hartmut Michel*

* Max-Planck-Institut für Biophysik, Heinrich-Hoffmann-Strasse 7, D-60528, Frankfurt am Main, Germany

† Johann Wolfgang Goethe-Universität, Institut für Mikrobiologie, Marie-Curie-Strasse 9, D-60439, Frankfurt am Main, Germany

‡ Present address: Skirball Institute of Biomolecular Medicine, NYU Medical Center, 540 First Avenue, New York, New York 10016, USA

Fumarate reductase couples the reduction of fumarate to succinate to the oxidation of quinol to quinone, in a reaction opposite to that catalysed by the related complex II of the respiratory chain (succinate dehydrogenase). Here we describe the crystal structure at 2.2 Å resolution of the three protein subunits containing fumarate reductase from the anaerobic bacterium *Wolinella succinogenes*. Subunit A contains the site of fumarate reduction and a covalently bound flavin adenine dinucleotide prosthetic group. Subunit B contains three iron–sulphur centres. The menaquinol-oxidizing subunit C consists of five membrane-spanning, primarily helical segments and binds two haem *b* molecules. On the basis of the structure, we propose a pathway of electron transfer from the dihaem cytochrome *b* to the site of fumarate reduction and a mechanism of fumarate reduction. The relative orientations of the soluble and membrane-embedded subunits of succinate:quinone oxidoreductases appear to be unique.

Succinate dehydrogenases (succinate:quinone reductases, SQR) and fumarate reductases (quinol:fumarate reductases, QFR) catalyse the oxidation of succinate to fumarate as well as the reverse reaction. SQR (complex II) is involved in aerobic metabolism as part of the citric acid cycle and of the aerobic respiratory chain¹. QFR is involved in a form of anaerobic respiration with fumarate as the terminal electron acceptor^{2,3}, and is part of the electron transport chain catalysing the oxidation of various donor substrates (such as NADH, H₂ or formate) by fumarate. These reactions are coupled by

an electrochemical proton gradient to phosphorylation of ADP with inorganic phosphate by ATP synthase.

QFR and SQR complexes are collectively referred to as succinate:quinone oxidoreductases (EC 1.3.5.1) and are predicted to share similar structures. The complexes consist of two hydrophilic and one or two hydrophobic, membrane-integrated subunits⁴. The larger hydrophilic subunit A carries a covalently bound flavin adenine dinucleotide (FAD), and subunit B contains three iron–sulphur centres. QFR of *W. succinogenes* and SQR of *Bacillus subtilis*

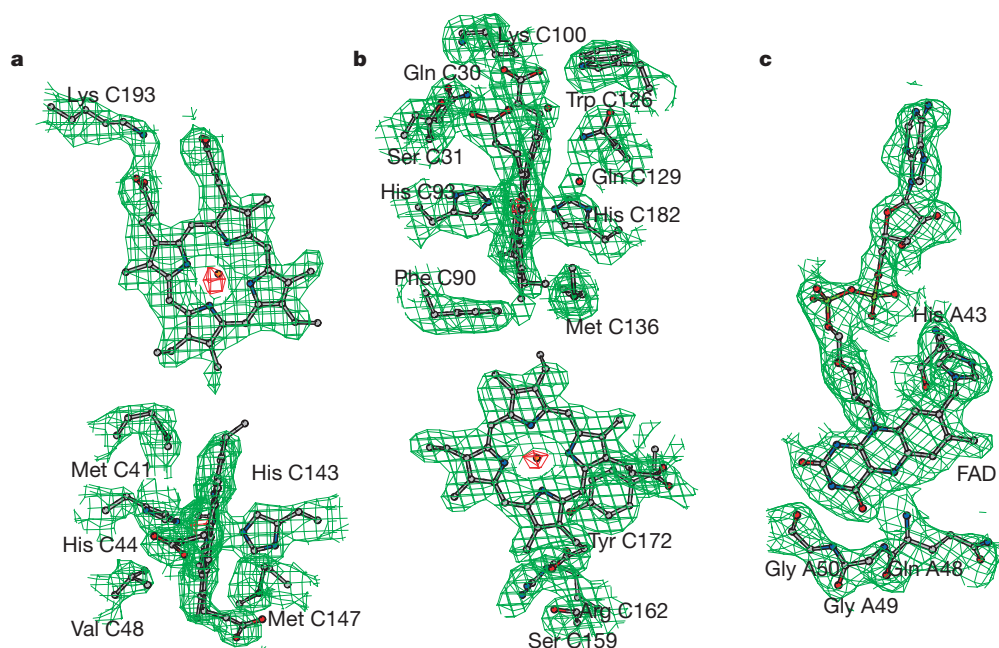


Figure 1 Representative parts of the experimental electron-density maps for crystal form A calculated with the MIRAS phases after density modification and phase extension to 2.2 Å resolution. C, N, O, P and S atoms are shown in grey, blue, red, light green and green, respectively; haem iron centres are shown in orange. Contour levels are 1.0 σ (green) and 9.0 σ (red) above the mean density of the map. Figs 1–4 and 6 were prepared

with a version of Molscript⁴⁶ modified by R. Esnouf for colour ramping⁴⁷ and map drawing⁴⁸ capabilities. **a, b**, The two haem *b* molecules (*b_P* in the top half; *b_D* in the bottom half of each panel) and the side chains of some neighbouring residues in the transmembrane region. **c**, The covalently bound FAD prosthetic group.

contain only one hydrophobic subunit (C) with two haem *b* groups. In contrast, SQR and QFR of *Escherichia coli* contain two hydrophobic subunits (C and D) which bind either one (SQR) or no haem *b* group (QFR). Recently, the structure of this haem-free *E. coli* QFR has been described at a resolution of 3.3 Å ($R_{\text{free}} = 29.2\%$; 49,332 unique reflections)⁵.

We have determined at 2.2 Å resolution the structure of the *W. succinogenes* QFR containing two haem groups by X-ray crystallography. Here we describe the overall structure of the complex, the structures of the three subunits and the arrangement and binding mode of the six prosthetic groups. As a result, we propose a pathway of electron transfer and a mechanism of fumarate reduction. We also compare the structure with that of the *E. coli* enzyme. The relative orientations of the soluble and the membrane-embedded subunits differ between the two enzymes.

Structure determination

The fully oxidized fumarate reductase from *W. succinogenes* was crystallized (see Methods). Under these conditions we obtained crystals with two different unit cells, both belonging to space group $P2_1$. Crystal form 'A' has unit cell dimensions $a = 85.2$ Å, $b = 189.0$ Å, $c = 117.9$ Å and $\beta = 104.5^\circ$. Crystal form 'B' has unit cell dimensions $a = 118.4$ Å, $b = 85.1$ Å, $c = 188.9$ Å and $\beta = 96.5^\circ$. In both cases, there are two fumarate reductase complexes in the asymmetric unit. Data collection statistics and the quality of initial phases determined for crystal form A are summarized in Table S1 (see Supplementary Information). The phases were improved using a series of density-modification methods. Representative electron-density maps are shown in Fig. 1. The excellent quality of the experimental map allowed rapid model building for most of the structure. We then used the partially refined structural model for crystal form A ($R = 26.5$, $R_{\text{free}} = 28\%$) as a search model for the phase determination of crystal form B by molecular replacement (see Supplementary Information). The refined structural models for crystal forms A ($R = 21.2\%$, $R_{\text{free}} = 22.4\%$; Table 1) and B ($R = 21.3\%$, $R_{\text{free}} = 22.3\%$) contain 2,296 of the 2,302

residues of the fumarate reductase dimer. Only the carboxy-terminal residue A656 of each A subunit and the two C-terminal residues C255 and C256 of each C subunit are not included. (Throughout this article, single letters in residue names refer to subunit identifiers, and amino-acid residues are represented by their three-letter codes.) For each asymmetric unit, two FAD groups, two Fe_2S_4 , two Fe_4S_4 and two Fe_3S_4 iron-sulphur centres, and four haem *b* molecules, are included in the current models. The structures contain two metal cations and two dodecyl β -D-maltoside detergent molecules as well as 691 (form A) and 504 water molecules (form B), respectively. In addition, the structure for crystal form B also contains two fumarate molecules. The ratios of the number of independent observations to the number of parameters required to define the model, $n_{\text{obs}}/n_{\text{par}}$ (Table 1), are 2.31 (form A) and 2.02 (form B). The corresponding value for the structure of *E. coli* fumarate reductase⁵ is 0.73.

The experimental electron density (Fig. 1a–c) was of sufficiently high quality to identify a region of nine residues (Asp A281 to Pro A289) where the previously determined⁶ sequence was incorrect. This has since been confirmed by resequencing (J. Simon and A.K., unpublished data).

Fumarate reductase architecture

In both crystal forms, the two fumarate reductase complexes in the asymmetric unit are associated in an identical fashion, forming a dimer. A view of the *W. succinogenes* fumarate reductase dimer parallel to the membrane is shown in Fig. 2a. In projection, as viewed in Fig. 2a, the part of the dimer that is integrated into the membrane has a trapezoidal appearance. It comprises ten membrane-spanning, primarily helical segments and their connections. These features of trapezoidal appearance and the smaller periplasmic surface as compared to the cytoplasmic surface resemble those of the structure of cytochrome *c* oxidase, complex IV of the respiratory chain⁷. The globular subunits B and A are attached to the trapezoid from the top. When viewed along the membrane normal, as in Fig. 2b, the membrane part of the fumarate reductase dimer is ellipsoidal.

About $3,665$ Å² (8%) of the monomer surface is buried upon dimer formation. This is more than eleven times the value of ~ 325 Å² reported⁵ for the *E. coli* fumarate reductase dimer, indicating that the dimer in the present structure is much more tightly associated. The portion of buried surface varies for the different subunits and is highest for the C subunit (12.2%). It is 2.6% for the A subunit and negligible for the B subunit. The dimer interface is mainly apolar, but 28% of the C subunit dimer interface is polar.

Protein subunits and prosthetic groups

The protein subunits are shown as $\text{C}\alpha$ traces in Fig. 3a–d.

Subunit A. This subunit has a relative molecular mass of 73,000 ($M_r = 73\text{K}$)⁶ and comprises four domains (Fig. 3a): a large FAD-binding domain (residues A1–260 and A366–436), a capping domain (A260–366), a helical domain (A436–554) consisting of a single helix and a three-helix bundle, and a C-terminal domain (A554–655), consisting of a pair of antiparallel β -strands followed by a longer and a short helix (Fig. 3a).

The FAD-binding domain with its Rossmann-type fold is structurally similar to some other FAD-binding proteins. The α -carbon atoms of the FAD-binding domain superimpose on those of thioredoxin reductase⁸ with a root mean square (r.m.s.) deviation of 3.4 Å for 211 of the thioredoxin reductase residues, as determined using DALI⁹.

Of the 655 $\text{C}\alpha$ atoms in the model of the A subunit, 358 (54.7%) could be aligned (with an r.m.s. deviation of 1.7 Å) to a model containing 478 residues of *E. coli* L-aspartate oxidase (PDB entry 1CHU, provided before release by A. Mattevi¹⁰), with 109 residues (16.6% of the A subunit) found to be identical. The degree of similarity was higher for the FAD-binding domain and the three-

Table 1 Refinement statistics

	Form A native	Form B native
Resolution range (Å)	38.63–2.20	38.87–2.33
No. of reflections used (completeness)	178,755 (98.0%)	152,939 (95.8%)
in working set	176,275 (96.6%)	150,440 (94.3%)
in test set	2,480 (1.4%)	2,499 (1.6%)
R_{free} (%) (38.63–2.33 Å)*	22.4 (22.0)	22.3
R_{cryst} (%) (38.63–2.33 Å)†	21.2 (20.8)	21.3
Luzzati coord. error (Å)‡	0.26 (0.28)	0.28 (0.30)
SIGMAA coord. error (Å)§	0.21 (0.21)	0.25 (0.25)
Highest resolution shell (Å)	2.25–2.20	2.39–2.33
$I/\sigma(I)$	2.9	2.8
Completeness	99.2%	99.7%
No. of non-hydrogen atoms in the model		
Atoms with occupancy >0	19,055	18,632
Average B-factor (Å ²)	32.5	39.0
No. of solvent molecules	691	504
$n_{\text{obs}}/n_{\text{par}} $	2.31	2.02
r.m.s. deviations from ideal values¶		
bond lengths (Å)	0.007	0.007
bond angles (°)	1.301	1.298
dihedral angles (°)	21.7	21.5
improper angles (°)	1.64	1.64
NCS r.m.s. difference (Å)	0.015	0.009

* R_{free} is the free R -value⁴⁰ calculated for the full resolution range and for the 38.63 to 2.33 Å range (in parentheses) = $\sum_{hkl \in T} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl \in T} |F_{\text{obs}}|$, where T is the test set (one reflection in each of at least 2,400 thin resolution shells⁴¹ of the data).

† R is the crystallographic R -factor, calculated for the full resolution range of the working set and for the 38.63 to 2.33 Å range (in parentheses) = $\sum_{hkl \in T} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl \in T} |F_{\text{obs}}|$.

‡ Estimate of the mean coordinate error from a Luzzati plot⁴².

§ Estimate of the mean coordinate error from a SIGMAA plot⁴³.

|| n_{obs} , number of observed unique reflections used in the working set; n_{par} , number of parameters necessary to define the model; this includes three parameters (x, y, z coordinates) per atom plus one (isotropic atomic B factor).

¶ Based on protein parameters files⁴⁴, haem cofactor parameter files⁴⁵ and parameter files generated for the other prosthetic groups and the substrate (see Supplementary Information).

helix bundle of the helical domain and lower for the capping domain and the C-terminal domain. Subunit A of the structure of *E. coli* fumarate reductase can be superimposed on subunit A of the present structure with an r.m.s. deviation of 1.3 Å for 531 C α atoms, illustrating the high degree of structural similarity for the first three domains of this subunit.

FAD binding. The A subunit of all described membrane-bound succinate:quinone oxidoreductase complexes contains a covalently

bound FAD prosthetic group^{11,12}. In *W. succinogenes* fumarate reductase, His A43 binds FAD by an 8 α -[N ϵ -histidyl]-linkage to the isoalloxazine ring¹³. FAD is apparently further held in place by hydrogen bonds donated by the peptide NH groups of eleven residues (A13, A15, A16, A36, A44, A49, A50, A181, A393, A409 and A410) and the O γ atoms of four Ser residues (A35, A42, A44 and A409). With the exception of residue A44, which is a Thr in most other succinate:quinone oxidoreductases, these Ser residues are highly conserved. The peptide CO group of residue A50 accepts a hydrogen bond from FAD (Fig. 1c). The flavin is also indirectly hydrogen-bonded through water molecules to the peptide NH groups of five residues (A14, A17, A43, A48 and A394) and to the peptide CO groups of residues A215, A216 and A319.

A new metal-binding site. During refinement, we identified an apparent additional metal-binding site. It is located in the FAD-binding domain of subunit A, less than 8 Å from the FAD group and less than 12 Å from its isoalloxazine ring system. Five backbone carbonyl oxygens (A371, A372, A373, A393, A395) as well as one tightly bound water molecule are involved in the binding of this metal cation of unknown function. In the refined models, it has been tentatively assigned to be a Ca²⁺ ion.

The dicarboxylate site. The data set of crystal form B is based on a crystal grown in the presence of 1 mM fumarate instead of 1 mM malonate. In the refined structure, we found significant difference density corresponding in shape to a fumarate molecule close to the isoalloxazine ring of FAD (Fig. 4a). Apparently, the dicarboxylate-binding site is mainly formed by the isoalloxazine ring of FAD, two arginine side chains (Arg A301 and Arg A404), one histidine side chain (His A369; Fig. 4b) and the side chain of Phe A141 (not shown in Fig. 4). One carboxylate group of fumarate is bound by polar interactions with Arg A404 and His A369, the other by Arg A301.

As viewed in Fig. 4, the left carboxyl group of the fumarate molecule is distorted. Its oxygen atoms are out of plane with respect to the rest of the fumarate molecule. Quantum chemical calculations at the MP2/6-31G* level (M. Hutter & C.R.D.L., unpublished data) indicate that the energy required to distort the molecule is 4.6 kcal mol⁻¹. This could be accounted for by one additional hydrogen bond, which can be donated by His A369. For an idealized planar fumarate molecule, the closest carboxyl oxygen would be ~0.9 Å further away from this residue. In addition, based on the position of the rest of the molecule, binding of an idealized planar fumarate would not be favoured owing to unfavourable steric interactions of the left carboxyl group with the FAD isoalloxazine ring.

If the imidazole ring of His A257 (not shown) were to be flipped by 180° around the χ_2 torsional angle, its N ϵ atom would be within hydrogen-bonding distance of the fumarate γ -carboxylate. However, this flip would also lead to possibly unfavourable interactions of the N ϵ atom with the guanidino group of Arg A301 and of the N δ atom with the peptide NH of Thr A259. Also, the hydrogen bond donated by the N ϵ atom to the O γ of Thr A259 would be lost. Therefore, the orientation of His A257 as included in the model is determined by its arrangement within the conserved 'HPT triad' of His A257, Pro A258 and Thr A259.

Tightly bound by hydrogen bonds from both Arg A301 and Arg A404, there is a water molecule, which is also within hydrogen-bonding distance of the fumarate α -methenyl C atom. Within hydrogen-bonding distance of the fumarate β -methenyl C atom, there is a second water molecule (Fig. 4a) which is also within hydrogen-bonding distance of N5 of the FAD isoalloxazine ring and the peptide NH of A48. The side chain O γ atom of Ser A409 is in a position to form hydrogen bonds with the N1 and O2 atoms of the FAD isoalloxazine ring and the guanidino group of Arg A404. All residues shown in Fig. 4 are widely conserved among succinate:quinone oxidoreductases.

Whereas the proximity of two Arg side chains and of FAD to the dicarboxylate binding site was expected¹⁴, the location of the Arg side chains in the sequence and the presence of His A369 at the

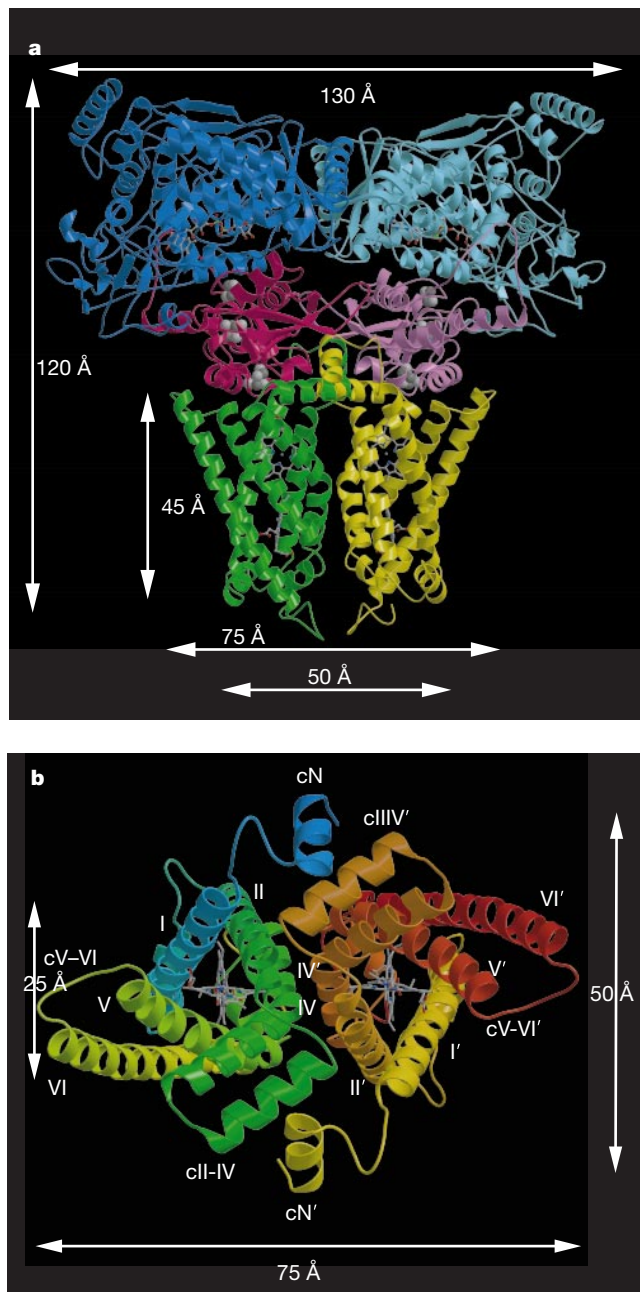


Figure 2 The three-dimensional structure of fumarate reductase. **a**, The fumarate reductase dimer viewed parallel to the membrane. The polypeptide backbones of the two A subunits are shown in blue and light blue, those of the two B subunits in red and pink, and those of the C subunits in green and yellow. Subunit A contains a covalently bound FAD. Subunit B contains three iron–sulphur clusters (Fe₂S₂, Fe₄S₄ and Fe₃S₄). The membrane-embedded subunit C contains two haem *b* molecules. **b**, View of the transmembrane helices of the subunit C dimer along the membrane normal from the cytoplasmic side. One monomer is colour-coded from blue (N terminus) to yellow (C terminus), the other from yellow (N terminus) to red (C terminus). The transmembrane helices are labelled I, II, IV, V and VI (ref. 22). Secondary structures were assigned using DSSP⁴⁹. Figure rendered with Raster3D⁵⁰.

binding site were less obvious. Particularly unexpected is the finding that Arg A273, which is homologous to Arg A248 of *E. coli* fumarate reductase, is absent from the inner shell of residues at the active site. For the *E. coli* enzyme, this residue was deduced by site-directed mutagenesis to be directly involved in substrate binding¹⁵. In the

W. succinogenes enzyme, Arg A273 is located in a segment containing four conserved Gly residues (A271, A274, A276 and A277). In the structure, the Arg side chain is oriented away from the dicarboxylate site. Mutation of this residue apparently has larger structural consequences.

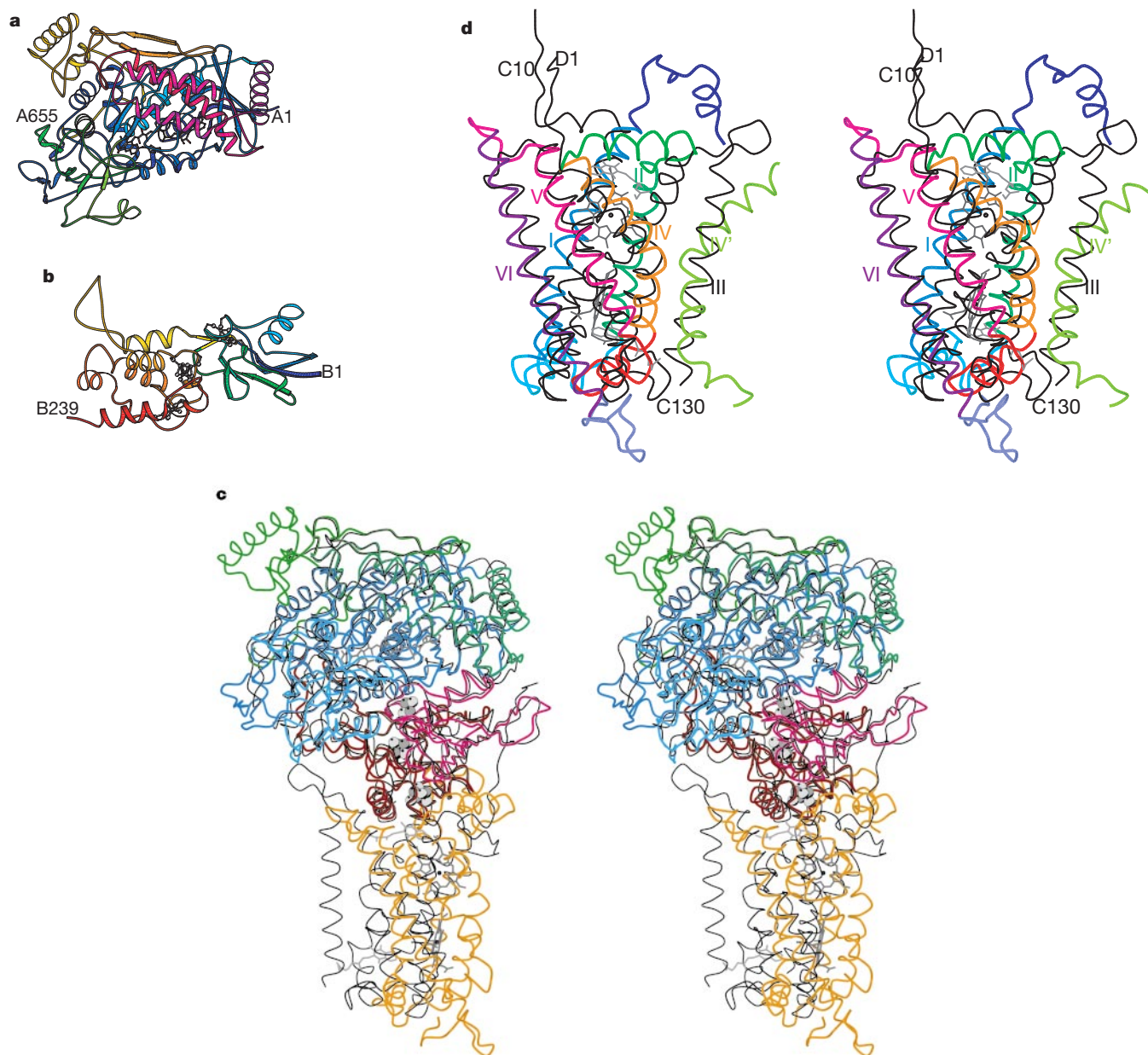


Figure 3 Structure of the subunits of *W. succinogenes* fumarate reductase. The structure of *E. coli* fumarate reductase⁵ is shown in black in **c** and **d** for comparison. The prosthetic groups FAD, haem *b_P* and haem *b_D* of the *W. succinogenes* enzyme are drawn as stick models and the iron–sulphur centres as grey spheres. The quinone models of the *E. coli* enzyme are in grey. **a**, *W. succinogenes* subunit A. Residues A1–260 and A366–436 constitute the FAD-binding domain (colour-coded from dark blue to light blue by sequence); A260–366 form the capping domain (dark green to light green); and A436–554 constitute the helical domain (purple to red) with a single helix (A437–456) followed by a three-helix bundle (A463–477, A484–501, A515–539). The C-terminal domain (A554–655, orange to yellow) includes an antiparallel pair of β -strands (A562–567, A577–582) followed by a longer (A607–625) and a short (A730–637) helix. **b**, *W. succinogenes* subunit B. The N-terminal plant ferredoxin domain (B1–98, colour-coded from blue (N terminus) to green (C-terminal end of domain)) consists of a pair of antiparallel β -strands (B4–11 and B22–29), followed by a helix (B36–46) which precedes the Cys residues (B57, B62, B65 and B77), which are ligands to the Fe₂S₂ iron–

sulphur centre. The C-terminal bacterial ferredoxin domain (B98–239, yellow (N-terminal) to red (C-terminus)) consists of five helices, two preceding (B108–117 and B139–148) the first cluster of Cys residues (B151, B154, B157 and B161) and another two preceding the second cluster of Cys residues (B208, B214 and B218), followed by a C-terminal helix (B224–235). **c**, Stereo view. *W. succinogenes* subunit A domains in blue (FAD-binding domain), light blue (capping domain), blue-green (helical domain) and green (C-terminal domain) as in **a**. Subunit B domains in pink (plant ferredoxin domain) and brown (bacterial ferredoxin domain) as in **b**. Subunit C is orange. **d**, Stereo view. *W. succinogenes* subunit C consists of five transmembrane helices, two periplasmic and two cytoplasmic helices. The N-terminal cytoplasmic helix (dark blue) is followed by transmembrane helix (TH) I (blue), a short periplasmic helix (light blue), TH II (blue-green), a second cytoplasmic helix (green), TH IV²² (orange), which is strongly kinked at position C137, a second periplasmic helix (red), TH V (pink) and TH VI (purple). Haem *b_P* (top half) and haem *b_D* (bottom half) are dark grey.

The side chain of residue Cys A272, proposed⁶ to be the site of inhibition by sulphhydryl reagents¹⁶ such as pCMS (4-chloromercuriphenylsulphonate), is disordered and was not found to be labelled by pCMB (see Supplementary Information) or pCMS (C.R.D.L. & H.M., unpublished data). Instead, these reagents labelled residue Cys A6 which, in the 22 most related sequences, is only conserved in the fumarate reductase of *Helicobacter pylori*¹⁷ and in the L-aspartate oxidase of *E. coli*¹⁸.

Chemistry of fumarate reduction. The arrangement of residues forming the active site of the oxidized enzyme with fumarate (Fig. 4a) suggests a mechanism of fumarate reduction as outlined in Fig. 5: FADH₂ presumably donates a hydride ion from its N5 to the β -methenyl group of fumarate. This hydride transfer could occur directly (Fig. 5a). In synchrony with this process, two of the double-bond electrons of fumarate move toward the α -carbon, and not the positively charged imidazolium group of His A257 (as proposed in ref. 19), but the tightly-bound water molecule donates a proton to the α -methenyl group to stabilize the reduced substrate. This proton could then be replenished by transfer via Ser A409 and Arg A404 of the proton released at the N1 position upon oxidation of FADH₂ (Fig. 5b). The water molecule within hydrogen-bonding distance of the FAD N5 and the Ser O γ atom within hydrogen-bonding distance of the FAD N1 are likely to be involved in the reorientation of hydrogen bonds upon oxidation of FADH₂. Release of the product could be facilitated by domain movement of the capping domain, thus moving residue Arg A301 away from the dicarboxylate site (Fig. 4b).

Subunit B. This subunit ($M_r = 27K$)⁶ consists of two domains, an amino-terminal 'plant ferredoxin' domain (B1–98, Fig. 3b), bind-

ing the Fe₂S₂ iron–sulphur centre, and a C-terminal 'bacterial ferredoxin' domain (B98–239) binding the Fe₄S₄ and the Fe₃S₄ iron–sulphur centres, as was expected from sequence comparisons⁶. The α -carbon atoms of the plant ferredoxin domain superimpose on Fe₂S₂-ferredoxin²⁰ with an r.m.s. difference of 2.0 Å for 64 of the 96 Fe₂S₂-ferredoxin residues.

Although the homology to bacterial Fe₄S₄-ferredoxins of the sequence in the ligand-binding segments of the bacterial ferredoxin domain has been pointed out^{6,4}, no superposition of coordinates with DALI was possible. The fragments containing ligands of the Fe₃S₄ and Fe₄S₄ centres are inverted in sequence with respect to bacterial ferredoxins containing Fe₃S₄ and Fe₄S₄ centres. With an r.m.s. deviation of 1.4 Å, 222 C α atoms of subunit B from *E. coli* fumarate reductase⁵ can be superimposed onto those of the present coordinates (Fig. 3c), indicating that this subunit is structurally very similar in both enzymes.

More than one-third of the surface area of subunit B is buried upon formation of the fumarate reductase ABC monomer, with 22% of the surface buried by subunit A and 12% of the subunit B surface buried by subunit C.

Iron–sulphur centres. The Fe₂S₂ iron–sulphur centre is coordinated by Cys residues B57, B62, B65 and B77, as proposed on the basis of sequence alignments⁶. All four Cys residues are within segments that are in contact with the A subunit. For instance, the peptide CO of Cys B62 is hydrogen-bonded to the guanidino group of Arg A41, which is close to the point of covalent FAD attachment. The Fe₄S₄ iron–sulphur centre is ligated to the protein through Cys residues B151, B154, B157 and B218, and the Fe₃S₄ centre is coordinated by Cys residues B161, B208 and B214 (Fig. 3b). The last three residues

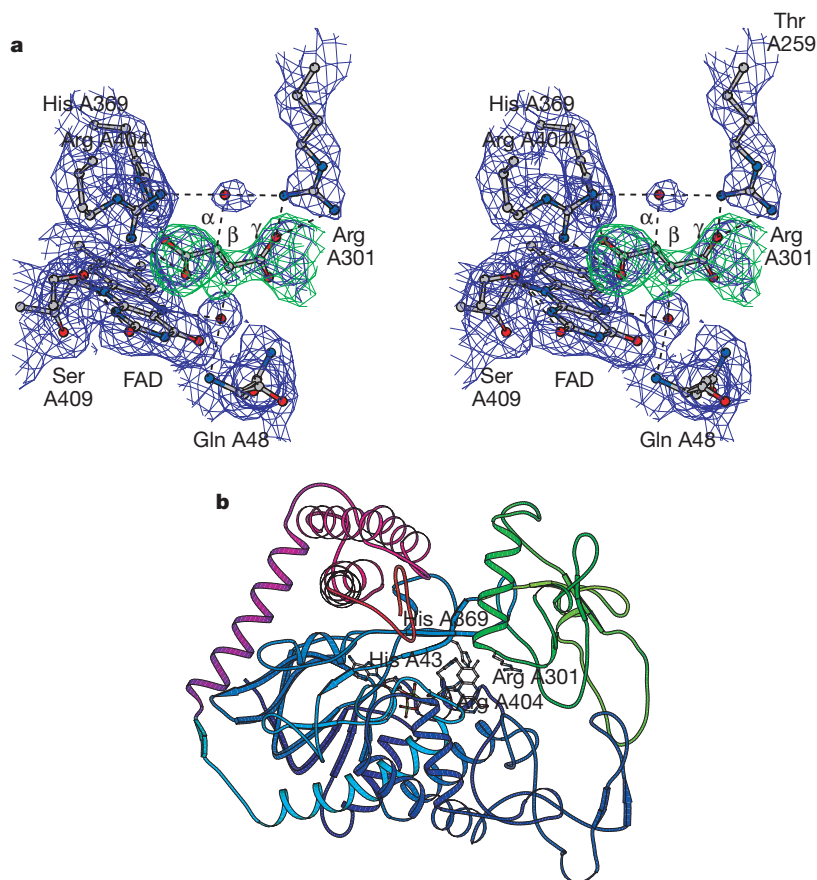


Figure 4 The dicarboxylate binding site in subunit A. **a**, Stereo view. SIGMAA-weighted electron-density maps from the refined model of crystal form B at 2.33 Å resolution. Contour levels are 1.0 σ ($2mF_o - DF_o$, purple) and 3.0 σ ($mF_o - DF_o$, green, with the fumarate molecule omitted from the phase calculation). Distances compatible with

hydrogen-bonding and salt-bridge interactions are indicated by dashed lines. Also labelled are the α - and β -methenyl and γ -carboxylate groups of fumarate (see text). **b**, Alternative view of the first three domains of subunit A, colour coded as in Fig. 3a, with the three residues Arg A301, His A369 and Arg A404 lining the binding site shown.

are within segments that are in contact with the C subunit.

Subunit C. This subunit ($M_r = 30\text{K}$)²¹ contains five membrane-spanning segments with preferentially helical secondary structures (Fig. 3d). These segments are labelled (according to ref. 22) I (C22–52), II (C77–100), IV (C121–149), V (C169–194) and VI (C202–237). According to the alignment of ref. 22, there is no transmembrane segment III in *W. succinogenes* fumarate reductase. To a varying degree, all five transmembrane segments are tilted with respect to the membrane normal. When viewed from the cytoplasmic side (Fig. 2b), the first four transmembrane segments appear to form a pore-like structure. The ‘pore’ is blocked by haems b_p and b_D (Fig. 2b). This is different from the haem arrangement in cytochrome *c* oxidase, where the membrane-bound haem molecules block two different pores⁷. The membrane-spanning segments are connected by four loops referred to as C-subunit loops pI–II, cII–IV, pIV–V and cV–VI, with c and p denoting loops on the cytoplasmic and periplasmic side of the membrane, respectively. Three of these loops contain short helices (pI–II: C55–63; cII–IV: C105–118; pIV–V: C158–164). The N terminus of subunit C is on the cytoplasmic side and the C terminus on the periplasmic side of the membrane. The N-terminal residues C3–11 form a helix denoted ‘cN’. The C-terminal residues C237–254 are of extended secondary structure with turns, and form the bottom of the C subunit. Residues C255 and C256 are disordered and are not visible in the electron-density map.

About one-tenth (9.7%) of the surface of subunit C is buried by forming contacts with subunit B. About one-eighth (12.2%) of the surface of subunit C is buried by interactions with the second subunit C molecule in the fumarate reductase dimer.

Haems b_p and b_D . Unlike the fumarate reductase of *E. coli*, which does not contain any haem groups, and the succinate dehydrogenase of *E. coli*, which contains one haem *b*, the C subunit of the *W. succinogenes* enzyme is a dihaem cytochrome b^{21} . The planes of both haem molecules are approximately perpendicular to the membrane surface and their interplanar angle is 95° (Figs 1, 2b).

The axial ligands to the ‘proximal’ haem b_p are His C93 of transmembrane segment II and His C182 of transmembrane segment V (Fig. 1b). This causes haem b_p to be located towards the cytoplasmic surface of the membrane, and thus towards the Fe_3S_4 iron–sulphur centre. Hydrogen bonds and salt bridges with the propionate groups of haem b_p are formed with the side chains of residues Gln C30, Ser C31, Lys C100, Trp C126 (Fig. 1b) and Lys C193 (Fig. 1a). Thus, side chains from the residues of all four transmembrane segments forming the pore described above are involved in the binding of haem b_p , underscoring the structural importance of the bound haem²³.

The axial ligands to the ‘distal’ haem b_D are His C44 of transmembrane segment I and His C143 of transmembrane segment IV (Fig. 1a), showing that all four haem axial ligands had been correctly predicted by sequence alignment²¹ and site-directed mutagenesis²³. Hydrogen bonds to the ring A propionate can be donated by the side chains of Ser C159 and Tyr C172, and a salt bridge with the Arg C162 side chain is possible (Fig. 1b). No such interactions can be found for the ring D propionate of haem b_D .

The binding of the two haem *b* molecules described here is very different from that described for the cytochrome bc_1 complex²⁴. In *W. succinogenes* fumarate reductase, the axial ligands for haem binding are located on four different transmembrane segments. In the cytochrome bc_1 complex, only two transmembrane segments are involved, each providing two axial haem *b* ligands. One consequence of this difference is that the distance between the two haem iron centres is distinctly shorter in fumarate reductase (15.6 Å) than it is in the cytochrome bc_1 complex (21 Å).

Relative orientation of soluble and membrane-embedded subunits. As discussed above for the individual soluble subunits, the structure of *E. coli* fumarate reductase can be superimposed on the present structure with an r.m.s. deviation of 1.4 Å for 757 C α atoms

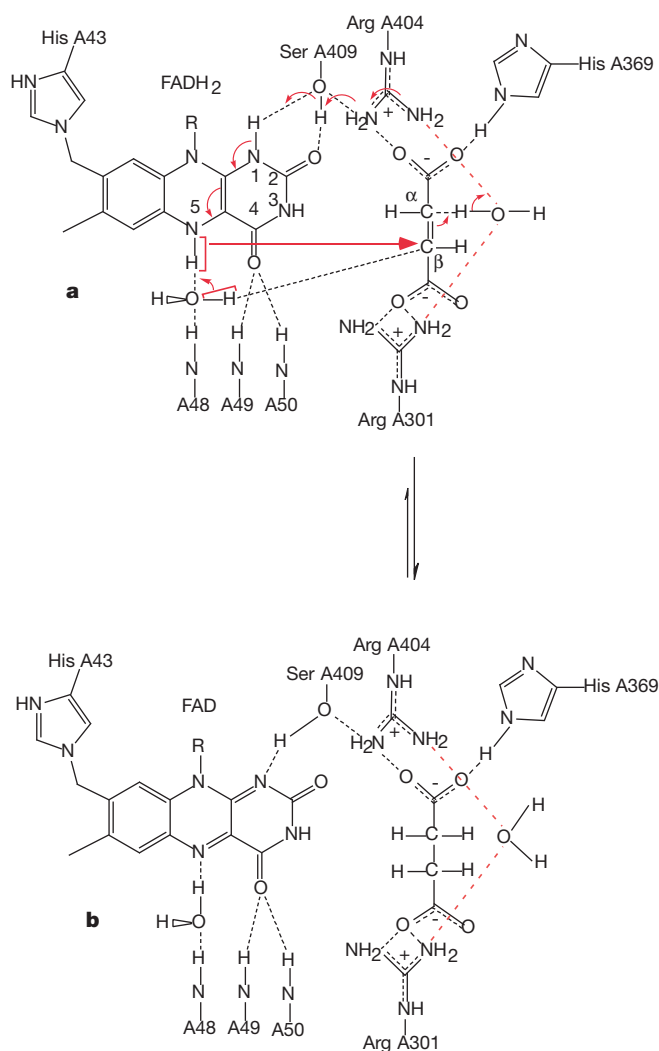


Figure 5 Possible mechanism of fumarate reduction in *W. succinogenes* fumarate reductase involving the residues shown in Fig. 4. In addition, the peptide backbone NH groups of residues A48–50 (Fig. 1c) are indicated. For clarity, only the isoalloxazine ring portion of FAD(H₂) is shown. Direct hydride ion transfer from the N5 of the FADH₂ isoalloxazine ring system (see Fig. 4) to fumarate is accompanied by proton transfer from the FADH₂ N1 position via a chain of hydrogen bonds, thus forming succinate.

from subunits A and B (Fig. 3c). This similarity in structure was expected on the basis of sequence comparisons. However, in this superimposition, the membrane-embedded subunits cannot be aligned. In an alternate superimposition, transmembrane subunits C and D of the *E. coli* enzyme can be overlaid onto the *W. succinogenes* C subunit with an r.m.s. deviation of 2.25 Å for 113 C α atoms from transmembrane helices I, II, IV, V and VI, and from the two periplasmic helices pI–II and pIV–V (Fig. 3d). Compared to the former superimposition (Fig. 3c), the latter (Fig. 3d) involves a rotation around the membrane normal of approximately 180° and a 25° rotation in the plane of the paper. This leads to two important conclusions. First, the structures of the transmembrane subunits carrying no haems and two haems, respectively, can be aligned to a significant degree, although only eleven of the aligned residues are identical. Second, the relative orientation of the soluble subunits and the transmembrane subunits is different in the fumarate reductase complexes from the two species.

The superimposition shown in Fig. 3d can be extended to 172 C α positions with a higher r.m.s. deviation of 3.0 Å. This increase is because the four transmembrane helices forming the pore shown in Fig. 2b are closer together in the haem-less *E. coli* enzyme than in

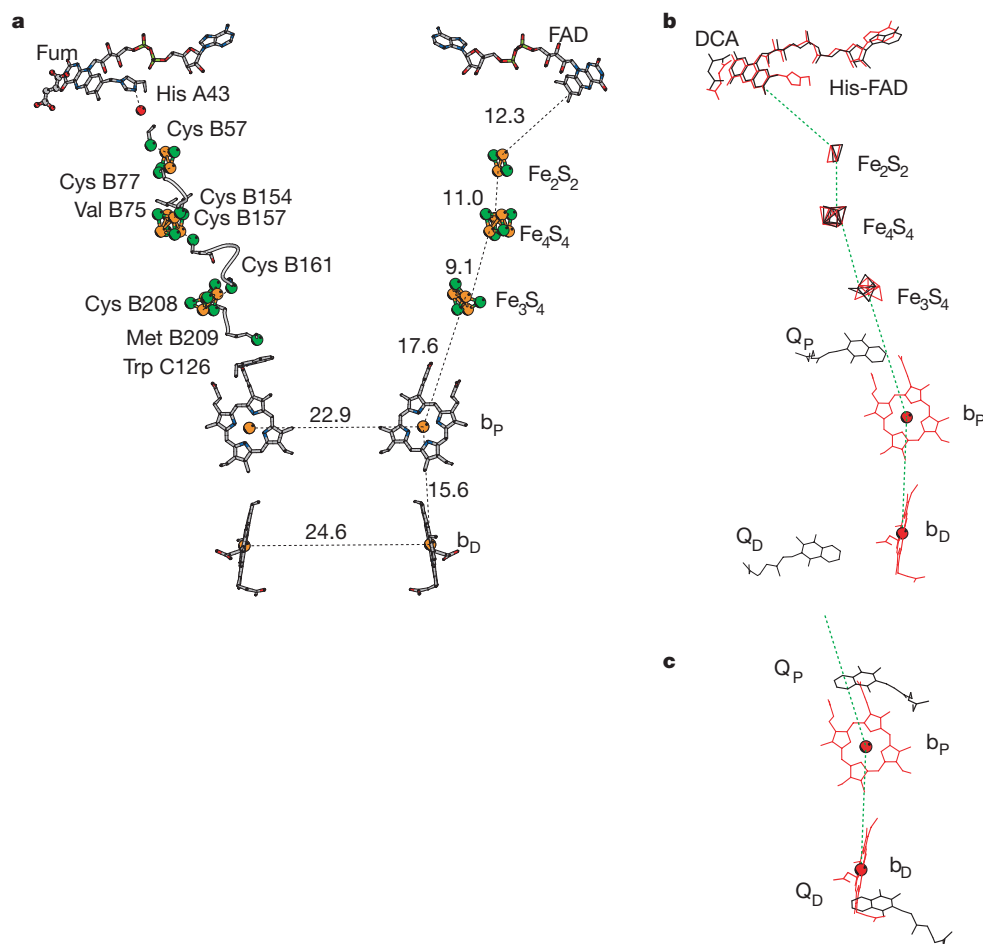


Figure 6 Arrangement of the prosthetic groups in the fumarate reductase dimer and possible electron-transfer pathways, including the position of bound fumarate (Fum). **a**, view as in Fig. 2a, colour coding as in Fig. 1. Numbers refer to indicated distances in Å. The side chain of Trp C126 is in a position to donate its hydrogen bond to the ring A propionate of haem b_P . This Trp side chain is in van der Waals contact (3.9 Å) with the S8 of Met B209, which in turn is connected by the polypeptide backbone with Cys B208, a ligand to the Fe_3S_4 iron–sulphur centre. Another ligand of the Fe_3S_4 centre, Cys B161, is connected, both by the polypeptide backbone and by a hydrogen bond, to Cys B157, a ligand of the Fe_4S_4 centre. Similarly, the third ligand of the Fe_3S_4 centre, Cys B214, is connected to the second ligand of the Fe_4S_4 centre, Cys B218 (not shown). A third ligand

of the Fe_4S_4 centre, Cys B154, is in van der Waals contact with Leu B75, which in turn is connected by the polypeptide backbone to Cys B77, a ligand to the Fe_2S_2 centre. A second ligand of the Fe_2S_2 centre is in van der Waals contact with a water molecule which is hydrogen-bonded to the N8 atom of His A43, the site of covalent attachment of FAD. **b, c**, Comparison of the electron-transfer pathways in the fumarate reductase complexes of *W. succinogenes* (red) and *E. coli* (black). Q_P , proximal quinone; Q_D , distal quinone. **b**, Structures are superimposed based on the Cα atoms of the A and B subunits (Fig. 3c). DCA, dicarboxylate (fumarate in the case of the *W. succinogenes* QFR coordinates (red) and oxaloacetate for the *E. coli* QFR coordinates (black)). **c**, Structures are superimposed based on the Cα atoms of the transmembrane subunits (Fig. 3d).

W. succinogenes fumarate reductase. For example, the distances between the Cα atoms of the *W. succinogenes* haem axial ligands are 13.0 Å (His C44–His C143) and 12.7 Å (His C93–His C182) for haems b_D and b_P , respectively. The corresponding distances in the *E. coli* enzyme are 9.9 Å (Leu C42–Met D31) and 11.2 Å (His C82–Cys D77). Another feature compensating for the absence of haems in the *E. coli* enzyme is the presence of more bulky residues (such as Arg C28, Trp C86 and Arg D81) where the corresponding positions in the *W. succinogenes* enzyme are occupied by smaller residues (Gln C30, Ala C97 and Gly C186). The superposition shown in Fig. 3d also indicates that the position of transmembrane helix III in the *E. coli* enzyme is occupied by transmembrane helix IV of the other *W. succinogenes* fumarate reductase monomer in the dimer. As the transmembrane subunit provides the largest contact area for the formation of the *W. succinogenes* QFR dimer, this superposition also explains the different degree of dimerization and different orientation of the monomers in the enzyme dimers of the two species.

Electron-transfer pathways

For fumarate reductase to function, electrons have to be transferred from the quinol-oxidizing site in the membrane to the fumarate-

reducing site, protruding about 40 Å into the cytoplasm. For *W. succinogenes* fumarate reductase, the experimental data²⁵ are consistent with this electron transfer from the quinol to fumarate occurring through (at least) one haem b group. The shortest distance between haem b_P and haem b_D is 4.2 Å. This distance is even shorter than that between haem a and haem a_3 in cytochrome c oxidase⁷, indicating that both haem b groups may be involved in electron transfer. In the event of transmembrane electron transfer via the haems, and if a 'distal' and a 'proximal' quinone-binding site are present (see below), electrogenic proton transfer, similar to the cytochrome bc_1 complex, is possible. This possibility awaits experimental verification.

The arrangement of the prosthetic groups in the fumarate reductase dimer is shown in Fig. 6a. As shown in Fig. 4, the fumarate molecule is in van der Waals contact with the isoalloxazine ring of FAD. The connectivity pattern shown in Fig. 6a therefore provides one straightforward pathway by which electrons could be transferred efficiently from the dihaem cytochrome b to the dicarboxylate-binding site.

The Fe_4S_4 iron–sulphur centre has a very low potential ($E_m = -250$ mV) and has been suggested not to participate in

electron transfer⁴. However, the determined low potential may be an artefact due to anti-cooperative electrostatic interactions between the redox centres²⁶. The position of the Fe₄S₄ centre as revealed in the structure presented here is highly suggestive of a direct role in electron transfer from the Fe₃S₄ centre to the Fe₂S₂ centre.

The prosthetic groups of *W. succinogenes* fumarate reductase are compared with those of the *E. coli* enzyme on the basis of the superposition of the soluble subunits (Fig. 6b) and the superposition of the membrane-embedded subunits (Fig. 6c). Correlated with the similarity in structure of the soluble protein subunits (Fig. 3c, above), the histidyl-FAD group and the three iron-sulphur centres superimpose well, with the exception of those Fe₃S₄ sulphur atoms that are oriented towards the membrane-embedded subunit(s).

Quinone-binding sites. The existence of up to three quinone-binding sites in succinate:quinone oxidoreductases is being discussed⁴: a 'proximal' binding site, close to haem *b_p* (where present) and the Fe₃S₄ iron-sulphur centre, and a 'distal' binding site close to haem *b_D*. In some cases, the proximal binding site has been reported to contain an EPR-detectable semiquinone pair⁴. Although dimethylnaphthoquinone was present in the crystallization droplets, no density for a quinone could be found in either crystal form of the oxidized enzyme presented here. By analogy with the cytochrome *bc₁* complex²⁴, the quinone-binding site(s) is (are) expected to be located in the vicinity of haem *b_p* (and haem *b_D*). By analogy to the photosynthetic reaction centre²⁷, the quinone(s) is (are) expected to bind in the vicinity of the Trp 'belts' in the hydrophobic surface-to-polar transition zone of the membrane (not shown).

The *E. coli* fumarate reductase coordinate set 1FUM (ref. 5) contains models for two quinone molecules per ABCD monomer. Although some of the B values of the quinone ring atoms are larger than 100 Å², indicating that these quinone models may not be well defined, these models have been included in Fig. 6b and c for comparison. In the superposition of the transmembrane subunits (Fig. 6c), the proximal quinone clashes with the ring A haem *b_p* propionate (Fig. 1b). The distal quinone clashes with the A haem *b_D* propionate. In summary, this superposition indicates that the quinone-binding site(s) of the two enzymes are not simply superimposable. In the *W. succinogenes* enzyme, they may be close to the haem molecules and the Trp belts, as previously inferred by comparison with other quinone-binding membrane proteins above. Note that quinones are the products and not the substrates of fumarate reductases. Therefore, binding of quinones might lead to substrate inhibition and has to be avoided. With this argument, the two quinones in the *E. coli* enzyme would have to be considered as structural quinones, with possible electron-transfer activities at least for the proximal quinone, and the existence of an additional quinol-binding site is required.

It is unlikely that transmembrane electron transfer occurs in the *E. coli* enzyme, because of the large gap of 27 Å between the two quinones. The close proximity between the two haems of *W. succinogenes* fumarate reductase as described above, however, offers a much more straightforward possibility for transmembrane electron transfer.

Conclusions

Most members of the succinate:quinone oxidoreductase superfamily contain at least one haem bound to the membrane-embedded subunit(s)²⁸. The well refined structure of the fumarate reductase presented here is the first of a haem-containing succinate:quinone oxidoreductase. At 2.20 Å resolution, it is currently the highest resolution structure available of a respiratory membrane protein complex. In this structure, the arrangement of two haems in the transmembrane domain is revealed. This arrangement suggests an efficient pathway for transmembrane electron transfer. The high resolution (2.33 Å) of the structure with bound fumarate allows us

to propose a catalytic mechanism for the reduction of fumarate to succinate. As all of the identified residues are conserved, this mechanism is of general relevance to all succinate:quinone oxidoreductases. In summary, the structure provides a framework for understanding the mechanism of complex II of the respiratory chain and related proteins at an atomic level. □

Methods

Enzyme purification and crystallization

W. succinogenes was grown with formate and fumarate as described²⁹. Fumarate reductase was purified as described³⁰, except that the hydroxyapatite column was replaced by a DEAE CL-6B column and the detergent Triton X-100 was replaced by a mixture of dodecyl-β-D-maltoside/decyl-β-D-maltoside on a second DEAE CL-6B column. The enzyme was then subjected to isoelectric focusing. We used the preparative flat-bed electrofocusing procedure for the preparation of photosystem I (ref. 31) with the following modifications: the pH range of the ampholytes was 5–8 (servalyt 5–8), and dodecyl-β-D-maltoside (0.05%) and decyl-β-D-maltoside (0.20%) were added to the ampholine solution. The gel band containing fumarate reductase was cut out and suspended in a buffer (pH 7.3) containing 20 mM HEPES, 1 mM EDTA and 2 mM malonate. After sedimentation of the gel material by centrifugation, the ampholytes were removed by gel filtration on PD-10 columns (Pharmacia).

The fully oxidized fumarate reductase from *W. succinogenes* was crystallized in the presence of 1 mM K₃Fe(CN)₆, 0.05% dodecyl-β-D-maltoside, 0.20% decylmaltoside, 2.4% benzamide, 5% polyethylene glycol (PEG-3350), 75 mM NaCl, 5% dimethylformamide (DMF), 1 mM dimethylnaphthoquinone (DMN), 1 mM malonate and 10 mM HEPES/10 mM citrate, pH 6.4, at a protein concentration of 9.5 mg ml⁻¹ by vapour diffusion against a reservoir containing 150 mM NaCl, 10% PEG-3350 and 20 mM citrate, pH 5.6. Nucleation was enhanced by applying the microseeding technique³².

Diffraction data

X-ray data were collected at ESRF beamline BM14 (λ = 0.996 Å, T = 2–4 °C), Grenoble. Owing to their thickness, the redox state of the crystals could not be monitored *in situ*. Intensity data were obtained using a CCD detector (MAR). Additional data were collected in-house with a rotating anode X-ray generator as source (CuKα, λ = 1.5418 Å) and an R-Axis IIC imaging plate as detector. Only one crystal was required for each dataset listed in Table S1 (see Supplementary Information). Data were processed with the HKL programs DENZO and SCALEPACK³³.

Phasing

Heavy-atom sites were determined from difference Patterson maps. The relative positioning of sites between derivatives was determined by the difference Fourier technique, using series of programs in the CCP4 suite³⁴. MIRAS (multiple isomorphous replacement with anomalous scattering) phasing and refinement of the heavy-atom positions were initially done using SHARP³⁵. Iron positions were determined iteratively from a consistent set of peaks in the SHARP anomalous residual maps of the native and derivative data sets, which were absent in the isomorphous residual maps. After inclusion of all 22 Fe sites in the asymmetric unit for all data sets, phase calculation and refinement was completed with MLPHARE³⁶.

Density modification

Density modification and phase extension to 2.20 Å were performed with the program DM³⁷. Solvent flattening using a solvent content of 60%, histogram matching and averaging were applied using the initial non-crystallographic symmetry (NCS) operator determined from the heavy atom and iron positions. Phase extension was performed in resolution steps, starting with the low-resolution data (50–5 Å), where the figure of merit was high. The initial real-space R value was 0.258, but it was reduced to 0.116 after 100 cycles, and the NCS correlation between regions of related density increased from 0.676 to 0.919.

Model building and refinement

The atomic model of fumarate reductase was built using program O³⁸. Simulated annealing, followed by conventional positional refinement and restrained individual B-factor refinement, was performed using CNS³⁹. Initially, strict NCS constraints were applied between the two fumarate reductase monomers in the asymmetric unit. In later cycles of refinement these were replaced by NCS restraints with a weight of 300 kcal mol⁻¹ Å⁻².

Received 23 June; accepted 14 October 1999.

1. Saraste, M. Oxidative phosphorylation at the fin de siècle. *Science* **283**, 1488–1493 (1999).
2. Kröger, A. Fumarate as terminal acceptor of phosphorylative electron transport. *Biochim. Biophys. Acta* **505**, 129–145 (1978).
3. Kröger, A. *et al.* Bacterial fumarate respiration. *Arch. Microbiol.* **158**, 311–314 (1992).
4. Hägerhäll, C. Succinate:quinone oxidoreductases. Variations on a conserved theme. *Biochim. Biophys. Acta* **1320**, 107–141 (1997).
5. Iverson, T. M., Luna-Chavez, C., Cecchini, G. & Rees, D. C. Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science* **284**, 1961–1966 (1999).

6. Lauterbach, F. *et al.* The fumarate reductase operon of *Wolinella succinogenes*. Sequence and expression of the *frdA* and *frdB* genes. *Arch. Microbiol.* **154**, 386–393 (1990).
7. Iwata, S., Ostermeier, C., Ludwig, B. & Michel, H. Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature* **376**, 660–669 (1995).
8. Kuriyan, J. *et al.* Convergent evolution of similar function in two structurally divergent enzymes. *Nature* **352**, 172–174 (1991).
9. Holm, L. & Sander, C. Mapping the protein universe. *Science* **273**, 595–602 (1996).
10. Mattevi, A. *et al.* Structure of L-aspartate oxidase: implications for the succinate dehydrogenase/fumarate reductase oxidoreductase family. *Structure* **7**, 745–756 (1999).
11. Singer, T. P. & McIntire, W. S. Covalent attachment of flavin to flavoproteins: Occurrence, assay, and synthesis. *Methods Enzymol.* **106**, 369–378 (1984).
12. Cole, S. T., Condon, C., Lemire, B. D. & Weiner, J. H. Molecular biology, biochemistry and bioenergetics of fumarate reductase, a complex membrane-bound iron-sulfur flavoenzyme of *Escherichia coli*. *Biochim. Biophys. Acta* **811**, 381–403 (1985).
13. Kenny, W. C. & Kröger, A. The covalently bound flavin of *Vibrio succinogenes* succinate dehydrogenase. *FEBS Lett.* **73**, 239–243 (1977).
14. van Hellemond, J. J. & Tielens, A. G. M. Expression and functional properties of fumarate reductase. *Biochem. J.* **304**, 321–331 (1994).
15. Schröder, I. *et al.* Identification of active site residues of *Escherichia coli* fumarate reductase by site-directed mutagenesis. *J. Biol. Chem.* **266**, 13572–13579 (1991).
16. Unden, G. & Kröger, A. An essential sulfhydryl group at the substrate binding site of the fumarate reductase of *Vibrio succinogenes*. *FEBS Lett.* **117**, 323–326 (1980).
17. Tomb, J.-F. *et al.* The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547 (1997).
18. Flachmann, R. *et al.* Molecular biology of pyridine nucleotide biosynthesis in *Escherichia coli*. Cloning and characterization of quinolate synthesis genes *nadA* and *nadB*. *Eur. J. Biochem.* **175**, 221–228 (1988).
19. Vik, S. B. & Hatefi, Y. Possible occurrence and role of an essential histidyl residue in succinate dehydrogenase. *Proc. Natl Acad. Sci. USA* **78**, 6749–6753 (1981).
20. Tsukihara, T. *et al.* Structure of the [2Fe-2S] ferredoxin I from *Aphanotothea sacrum* at 2.2 Å resolution. *J. Mol. Biol.* **216**, 399–410 (1990).
21. Körtner *et al.* *Wolinella succinogenes* fumarate reductase contains a dihaem cytochrome *b*. *Mol. Microbiol.* **4**, 855–860 (1990).
22. Hagerhall, C. & Hederstedt, L. A structural model for the membrane-integral domain of succinate:quinone oxidoreductases. *FEBS Lett.* **389**, 25–31 (1996).
23. Simon, J. *et al.* Deletion and site-directed mutagenesis of the *Wolinella succinogenes* fumarate reductase operon. *Eur. J. Biochem.* **251**, 418–426 (1998).
24. Xia, D. *et al.* Crystal structure of the cytochrome *bc1* complex from bovine heart mitochondria. *Science* **277**, 60–66 (1997).
25. Unden, G., Albracht, S. P. J. & Kröger, A. Redox potentials and kinetic properties of fumarate reductase complex from *Vibrio succinogenes*. *Biochim. Biophys. Acta* **767**, 460–469 (1984).
26. Salerno, J. C. Electron transfer in succinate:ubiquinone reductase and quinol:fumarate reductase. *Biochem. Soc. Trans.* **19**, 599–605 (1991).
27. Lancaster, C. R. D. Quinone-binding sites in membrane proteins: what can we learn from the *Rhodospseudomonas viridis* reaction centre? *Biochem. Soc. Trans.* **27**, 591–596 (1999).
28. Hederstedt, L. Bioenergetics—Respiration without O₂. *Science* **284**, 1941–1942 (1999).
29. Brönder, M., Mell, H., Stupperich, E. & Kröger, A. Biosynthetic pathways of *Vibrio succinogenes* growing with fumarate as terminal electron acceptor and sole carbon source. *Arch. Microbiol.* **131**, 213–223 (1982).
30. Unden, G., Hackenberg, H. & Kröger, A. Isolation and functional aspects of the fumarate reductase involved in the phosphorylative electron transport of *Vibrio succinogenes*. *Biochim. Biophys. Acta* **591**, 275–288 (1980).
31. Tsiotis, G. in *A Practical Guide to Membrane Protein Purification* (eds von Jagow, G. & Schagger, H.) 149–159 (Academic, San Diego, 1994).
32. McPherson, A. *Crystallization of Biological Macromolecules* 304–305 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1999).
33. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
34. Collaborative Computation Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
35. De La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
36. Otwinowski, Z. in *Proc. CCP4 Study Weekend, 25–26 Jan 1991, Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–86 (SERC Daresbury Laboratory, Warrington, UK, 1991).
37. Cowtan, K. dm: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **30**, 34–38 (1994).
38. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors within these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
39. Brünger, A. T. *et al.* Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
40. Brünger, A. T. Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475 (1992).
41. Kleywegt, G. J. & Brünger, A. T. Checking your imagination: Applications of the free *R* value. *Structure* **4**, 897–904 (1996).
42. Luzzati, P. V. Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.* **5**, 802–810 (1952).
43. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).
44. Engh, R. H. & Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392–400 (1991).
45. Lancaster, C. R. D. & Michel, H. The coupling of light-induced electron transfer and proton uptake as derived from crystal structures of reaction centres from *Rhodospseudomonas viridis* modified at the binding site of the secondary quinone, Q_B. *Structure* **5**, 1339–1359 (1997).
46. Kraulis, P. J. MolScript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
47. Esnouf, R. M. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graphics Mod.* **15**, 132–134 (1997).
48. Esnouf, R. M. Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallogr. D* **55**, 938–940 (1999).
49. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
50. Merritt, E. A. & Bacon, D. J. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).

Supplementary information is available at Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank A. Mattevi and co-workers for providing the coordinates of the *E. coli* L-aspartate oxidase before release for the Protein Data Bank; S. Gemeinhardt, O. Schürmann (both Institut für Mikrobiologie), C. Weiss, B. Marx, C. Münke, C. Wardenberg and D. Vinzenz (all at MPI Biophysik) for technical assistance; A. Thompson (EMBL Grenoble), G. Leonard and E. Mitchell (both ESRF) for support during data acquisition and preliminary data processing at ESRF Grenoble beamline BM14; and L.-O. Essen for providing purified tetrakis-(acetoxymethyl)-methane (TAMM). This work was supported by the Deutsche Forschungsgemeinschaft (SFB 472, grants to A.K. and H.M.) and the Max-Planck-Gesellschaft.

Correspondence and requests for materials should be addressed to C.R.D. L. (e-mail: lancaster@mpibp-frankfurt.mpg.de) and H.M. (e-mail: michel@mpibp-frankfurt.mpg.de). Atomic coordinates of the structures for both crystal forms have been deposited in the Protein Data Bank (accession codes 1QLA (form A) and 1QLB (form B)).

Discovery of molecular hydrogen in a high-velocity cloud of the Galactic halo

P. Richter*, K. S. de Boer*, H. Widmann†, N. Kappelman†, W. Gringel†, M. Grewing† & J. Barnstedt†

* Sternwarte der Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

† Institute für Astronomie und Astrophysik, Universität Tübingen, Waldhäuserstrasse 64, D-72076 Tübingen, Germany

The Milky Way's halo contains clouds of neutral hydrogen with high radial velocities which do not follow the general rotational motion of the Galaxy¹. Few distances to these high-velocity clouds are known^{2,3}, so even gross properties such as total mass are hard to determine. As a consequence, there is no generally accepted theory regarding their origin. One idea^{4,5} is that they result from gas that has cooled after being ejected from the Galaxy through fountain-like flows powered by supernovae; another is that they are composed of gas, poor in heavy elements, which is falling onto the disk of the Milky Way from intergalactic space^{6,7}. The presence of molecular hydrogen, whose formation generally requires the presence of dust (and therefore gas, enriched in heavy elements), could help to distinguish between these possibilities. Here we report the discovery of molecular hydrogen absorption in a high-velocity cloud along the line of sight to the Large Magellanic Cloud. We also derive for the same cloud an iron abundance which is half of the solar value. From these data, we conclude that gas in this cloud originated in the disk of the Milky Way.

An earlier search for molecular gas in the Galactic halo was limited to radio observations of the carbon monoxide molecule (CO)^{8,9}. Molecular hydrogen (H₂) is by far the most abundant molecule in the interstellar medium, but in its cold form can only be observed in the far-ultraviolet (FUV) domain in absorption against bright background sources. Such measurements require a telescope working outside the Earth's atmosphere. The echelle spectrograph of the Orbiting and Retrieivable Far and Extreme Ultraviolet Spectrometer (ORFEUS)¹⁰ is the first instrument capable of investigating (in the FUV) the H₂ molecule outside the disk of the Milky Way. ORFEUS has improved wavelength coverage and better sensitivity than previous instruments.

For the investigation of the H₂ absorption in Galactic high-velocity clouds (HVCs), several lines of sight towards bright stars in the Magellanic Clouds are available from ORFEUS. The two Magellanic Clouds are the most nearby satellite galaxies of the Milky Way. The HVC in front of the 50-kpc-distant Large Magellanic Cloud (LMC) at a radial velocity (v_{rad}) of +120 km s⁻¹ is seen in almost all absorption spectra of LMC stars^{11,12}; this HVC is therefore known to have a substantial extent in the sky¹³, and therefore to be probably a foreground cloud of the Milky Way halo.

From the four ORFEUS spectra of LMC stars, the one to the B-type star HD269546 (Galactic longitude $l = 279.32^\circ$, latitude $b = -32.77^\circ$) shows the presence of H₂ at +120 km s⁻¹. This line of sight is the one in our sample with the highest column density of atomic hydrogen¹⁴. For all other LMC spectra no clear absorption is seen at HVC velocities, probably due to a combination of lower H I column densities, lower signal-to-noise ratios and line blending. Nine H₂ absorption lines from the lowest 7 rotational states and free from line blends were found in the spectrum of HD269546, and for 5 additional transitions upper limits for the equivalent widths were determined (Table 1). Figure 1 shows part of the spectrum where some of the detected H₂ absorption features are found. Figure 2 shows the Werner Q(2),0-0 line²⁴ at 1,010.941 Å, plotted on a velocity scale, as an example of the H₂ absorption pattern. The H₂

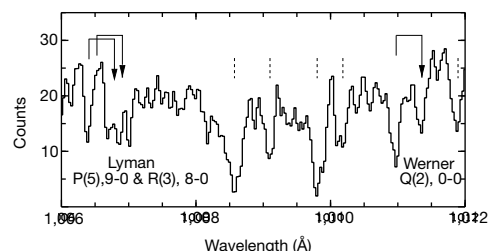


Figure 1 The spectral region between 1,006 and 1,012 Å in the ORFEUS spectrum of HD269546. Solid lines indicate H₂ absorption lines with the Galactic component at 0 km s⁻¹, and arrows show the HVC component at a Doppler velocity of about +120 km s⁻¹. Other identified absorption lines, which are of no importance, here, are marked with dashed lines.

absorption at $v_{\text{rad}} = +120$ km s⁻¹ matches perfectly the neutral-hydrogen emission profile of this sight line¹⁴, and there is no doubt that both neutral and molecular hydrogen belong to this HVC.

For the quantitative analysis, we made use of a standard curve-of-growth technique. We measured equivalent widths for the H₂ absorption at +120 km s⁻¹, and constructed curves of growth for each rotational state J individually which collectively fit the velocity-dispersion parameter $b = 9$ km s⁻¹. The high value of b indicates that the gas is turbulent. We find the H₂ column densities $N(J)$ to be as follows: $\log N(0) \leq 14.2$, $\log N(1) = 14.6 \pm 0.2$, $\log N(2) = 14.8 \pm 0.2$, $\log N(3) = 14.8 \pm 0.2$, $\log N(4) = 14.7 \pm 0.4$, $\log N(5) \leq 14.7$, $\log N(6) \leq 14.9$.

The total H₂ column density in this HVC is small, $N(\text{H}_2) = (2.2\text{--}3.6) \times 10^{15}$ cm⁻². Comparison with the column density of neutral hydrogen $N(\text{H I}) = 1.2 \times 10^{19}$ cm⁻², derived from the older Parkes 21-cm emission profile¹⁴, shows that the proportion of H₂ to H I is similar to that found in diffuse gas in the Galactic disc¹⁵. It is indicative of well illuminated gas with low dust content.

The distribution of column densities over the rotational states can be used to roughly estimate the excitation conditions in the H₂ gas¹⁶. The temperature of the H₂ gas is $\geq 1,000$ K, derived by a fit of the population of the rotational levels to a theoretical Boltzmann distribution. This value is probably not the actual kinetic temperature, but represents the excitation through ultraviolet photons from the Galactic disk. In sharp contrast, the relative abundance of *para*-H₂ (even J numbers) and *ortho*-H₂ (odd J numbers) for $J \leq 3$ is ≈ 1 . Thus the *ortho*-to-*para* ratio must have been thermalized by proton exchange¹⁷ with H⁺ to a gas temperature below 100 K. The discrepancy between these two temperature ranges is a sign that the gas is not in thermodynamical equilibrium. Further support for this is given by the fact that the observed low H₂ column density can hardly be described with approved theoretical models of H₂ formation and destruction processes¹⁸ when taking into account the strength of the radiation field from the Milky Way disk.

Table 1 H₂ equivalent widths for the HVC gas towards HD269546

H ₂ transition*	λ (Å)	$\log$ $f\lambda^\dagger$	W (mÅ)
LR(0), 4-0	1,049.366	1.39	≤ 30
LR(1), 8-0	1,002.457	1.26	34 ± 19
WQ(1), 1-0	986.798	1.56	75 ± 22
WQ(2), 0-0	1,010.941	1.38	65 ± 22
LR(2), 4-0	1,051.497	1.19	54 ± 23
LP(2), 7-0	1,016.472	1.01	≤ 57
LP(2), 11-0	975.343	0.81	≤ 42
LP(3), 3-0	1,070.142	0.89	33 ± 17
LR(3), 4-0	1,053.976	1.18	47 ± 23
LR(3), 8-0	1,006.418	1.19	63 ± 30
LR(4), 4-0	1,057.379	1.17	46 ± 23
LP(5), 4-0	1,065.594	1.00	≤ 31
LP(6), 3-0	1,085.382	0.94	≤ 54
LR(6), 7-0	1,030.064	1.17	65 ± 25

* Ref. 24.

† Ref. 23.

L, Lyman band; W, Werner band; λ , wavelength; W , equivalent width.

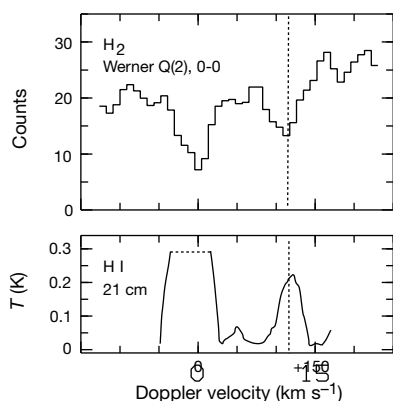


Figure 2 Comparison between the H_2 absorption profile (top panel) and the H I emission profile (bottom panel) for the HVC gas in the direction of HD269546. The H_2 $\text{Q}(2), 0-0$ line²⁴ at $1,010.941 \text{ \AA}$ in the Werner band (here plotted in the Doppler velocity scale, with respect to the local standard of rest) reveals H_2 absorption from HVC gas near $+120 \text{ km s}^{-1}$ (top panel). At this velocity, emission from neutral hydrogen¹⁴ (bottom panel) is also present.

Recent H I data from Parkes¹⁹ (Fig. 3) show that our sight-line crosses the HVC at $+120 \text{ km s}^{-1}$ in a region between two denser cores in a much larger complex. It shows spatial similarities to 'Chain A', recently found³ to be at a distance between 2.5 and 7 kpc. In view of the structure of the HVC, it seems plausible that the cloud core has recently fallen apart. If this is so, then cold H_2 from the cloud core was exposed to the unshielded ultraviolet flux from the Galactic disk that currently photodissociates the H_2 gas.

In order to further explore the origin of the HVC in front of the Magellanic Clouds, we have combined ORFEUS data with public data²⁵ from the International Ultraviolet Explorer (IUE) satellite and have determined the iron abundance in the HVC component at $+120 \text{ km s}^{-1}$. From an analysis of five Fe II absorption lines we derive $b = 9 \text{ km s}^{-1}$ and a column density of $\log N(\text{Fe II}) = 14.3 \pm 0.2$, in good agreement with an earlier investigation²⁰. This b -value is the same as for H_2 , supporting the idea that it is the same gas. The iron abundance is given by $\log N[\text{Fe II}/\text{H}] = -4.8$, which is half of the solar value²¹. This ratio is based on a value of $N(\text{H I})$ from the older $15'$ beam Parkes measurement¹⁴. In view of this high metallicity, we can rule out the possibility that the gas—and thus the H_2 —is of primordial origin or from intergalactic space²².

Our results favour the Galactic fountain model^{4,5} as explanation for the origin of the HVC gas in front of the LMC. In this model, hot gas is ejected out of the Galactic disk, cools down and falls back onto the Galactic plane. The molecular hydrogen must then have been formed within the cooled halo cloud, and at least a small amount of dust must be present in that HVC.

Theoretical considerations⁸ show that molecular hydrogen can form in HVCs once the hydrogen volume density $n(\text{H I})$ exceeds a critical value. This critical volume density depends on the H I column density, the amount of dust, the radiation field, the gas temperature and the velocity dispersion parameter b . We have found $b = 9 \text{ km s}^{-1}$ for our HVC gas, indicating that H_2 might form at volume densities as low as $n(\text{H I}) \approx 10 \text{ cm}^{-3}$. The actual volume density can only be determined by assuming a certain distance to the HVC in combination with assumptions about the depth structure of this cloud. However, the uncertainties of these parameters are too high to judge whether the theoretical models can describe the observations we have presented here. $\square$

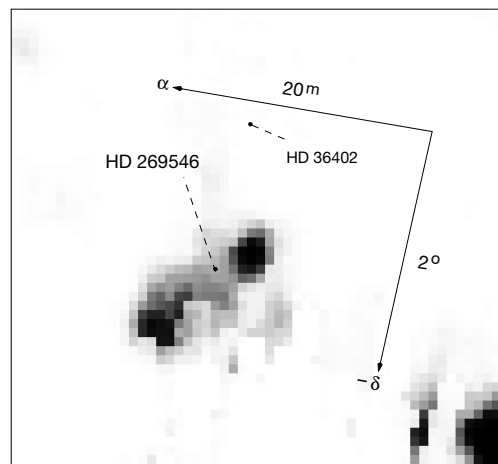


Figure 3 Map of the peak intensities for neutral hydrogen emission. The data were obtained from the HIPASS survey (preliminary, uncalibrated) for velocities $v_{\text{LSR}} = 105\text{--}145 \text{ km s}^{-1}$ in the field centred on HD269546 (equinox 2000; v_{LSR} , velocity with respect to the local standard of rest; α , right ascension; δ , declination). The map displays only the high intensities (dark structures). The line of sight towards HD269546 passes between two denser H I clumps in the HVC. The line of sight to the LMC star HD36402, on which absorption and emission is known at the same velocity¹¹, is also indicated. The map was made available by the HIPASS team¹⁹.

- van Woerden, H. *et al.* A confirmed location in the Galactic halo for the high-velocity cloud 'chain A'. *Nature* **400**, 138–141 (1999).
- Shapiro, P. R. & Field, G. B. Consequences of a new hot component of the interstellar medium. *Astrophys. J.* **205**, 762–765 (1976).
- Bregman, J. N. The galactic fountain of high-velocity clouds. *Astrophys. J.* **236**, 577–591 (1980).
- Oort, J. H. The formation of galaxies and the origin of the high-velocity hydrogen. *Astron. Astrophys.* **7**, 381–404 (1970).
- Blitz, L. *et al.* High-velocity clouds: building blocks of the Local Group. *Astrophys. J.* **514**, 818–843 (1999).
- Wakker, B. P., Murphy, E. M., van Woerden H. & Dame, T. M. A sensitive search for molecular gas in high-velocity clouds. *Astrophys. J.* **488**, 216–223 (1997).
- Akeson, R. L. & Blitz, L. A search for hydrogen and molecular gas absorption in high-velocity clouds. *Astrophys. J.* **523**, 163–170 (1999).
- Barnstedt, J. *et al.* The ORFEUS II Echelle Spectrometer: Instrument description, performance and data reduction. *Astron. Astrophys. Suppl.* **134**, 561–567 (1999).
- Savage, B. D. & de Boer, K. S. Ultraviolet absorption by interstellar gas at large distances from the Galactic plane. *Astrophys. J.* **243**, 460–484 (1981).
- Bomans, D. J., de Boer, K. S., Koornneef, J. & Grebel, E. K. CIV absorption from hot gas inside the supergiant shell LMC4 observed with HST and IUE. *Astron. Astrophys.* **313**, 101–112 (1996).
- de Boer, K. S., Morris, R. & Bajaja, E. The location of intermediate- and high-velocity gas in the general direction of the Large Magellanic Cloud. *Astron. Astrophys.* **233**, 523–526 (1990).
- McGee, R. X. & Newton, L. M. Neutral hydrogen in the Galactic halo. *Proc. Astron. Soc. Aust.* **6**, 358–385 (1986).
- Savage, B. D., Bohlin, R. C., Drake, J. F. & Budich, W. A survey of interstellar molecular hydrogen. *Astrophys. J.* **216**, 291–307 (1977).
- Spitzer, L. & Zweibel, E. G. On the theory of H_2 rotational excitation. *Astrophys. J.* **191**, L127–L130 (1974).
- Burton, M. G., Hollenbach, D. J. & Thielens, A. G. G. M. Mid-infrared rotational line emission from interstellar molecular hydrogen. *Astrophys. J.* **399**, 563–572 (1992).
- Jura, M. A. Interstellar clouds containing optically thin H_2 . *Astrophys. J.* **197**, 575–580 (1975).
- Putman, M. E. & Gibson, B. K. First results from the Parkes Multibeam High-velocity Cloud Survey. *Proc. Astron. Soc. Aust.* **16**, 70–76 (1999).
- Grewing, M. & Schulz-Lüpertz, E. Ionisationsstrukturen im galaktischen Halogas. *Mitt. Astron. Gesellsch.* **52**, 79–83 (1981).
- de Boer, K. S., Jura, M. A. & Schull, J. M. Diffuse and dark clouds in the interstellar medium. *Scientific Accomplishments of the IUE* (ed. Kondo, Y.) 485–515 (Reidel, Dordrecht, 1987).
- Pei, Y. C. & Fall, S. M. Cosmic chemical evolution. *Astrophys. J.* **454**, 69–76 (1995).
- Morton, D. C. & Dinerstein, H. L. Interstellar molecular hydrogen toward Zeta Puppis. *Astrophys. J.* **204**, 1–11 (1975).
- Field, G. B., Somerville, W. B. & Dressler, K. Hydrogen molecules in Astronomy. *Annu. Rev. Astron. Astrophys.* **4**, 207–244 (1966).
- Nichols, J. S. & Linsky, J. L. The Final Archive and recalibration of the International Ultraviolet Explorer (IUE) satellite. *Astron. J.* **111**, 517–536 (1996).

Acknowledgements

We thank M. Putman and C. Brüns for providing part of the HIPASS data.

Correspondence and requests for materials should be addressed to P.R. (e-mail: prichter@astro.uni-bonn.de).

Received 22 June; accepted 21 September 1999.

- Wakker, B. Distribution and origin of high-velocity clouds. *Astron. Astrophys.* **250**, 499–508 (1991).
- Wakker, B. P. & van Woerden, H. High-velocity clouds. *Annu. Rev. Astron. Astrophys.* **35**, 217–266 (1997).

Accretion of low-metallicity gas by the Milky Way

B. P. Wakker*, J. C. Howk†, B. D. Savage*, H. van Woerden‡, S. L. Tufte§, U. J. Schwarz||, R. Benjamin¶, R. J. Reynolds*, R. F. Peletier# & P. M. W. Kalberla*

* Department of Astronomy, † Department of Physics, University of Wisconsin, Madison, Wisconsin, 53706, USA
‡ Department of Physics & Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA
§ Kapteyn Institute, Postbus 800, 9700 AV Groningen, The Netherlands
¶ Department of Physics, Lewis & Clark College, Portland, Oregon 97219, USA
|| Astrophysics, University of Nijmegen, Postbus 9010, 6500 GL Nijmegen, The Netherlands
Department of Physics, University of Durham, Durham DH1 3LE, UK
* Radio-astronomisches Institut Universität Bonn, 53121 Bonn, Germany

Models of the chemical evolution of the Milky Way suggest that the observed abundances of elements heavier than helium ('metals') require a continuous infall of gas with metallicity (metal abundance) about 0.1 times the solar value. An infall rate integrated over the entire disk of the Milky Way of ~1 solar mass per year can solve the 'G-dwarf problem'—the observational fact that the metallicities of most long-lived stars near the Sun lie in a relatively narrow range^{1–3}. This infall dilutes the enrichment

arising from the production of heavy elements in stars, and thereby prevents the metallicity of the interstellar medium from increasing steadily with time. However, in other spiral galaxies, the low-metallicity gas needed to provide this infall has been observed only in associated dwarf galaxies⁴ and in the extreme outer disk of the Milky Way^{5,6}. In the distant Universe, low-metallicity hydrogen clouds (known as 'damped Ly α absorbers') are sometimes seen near galaxies^{7,8}. Here we report a metallicity of 0.09 times solar for a massive cloud that is falling into the disk of the Milky Way. The mass flow associated with this cloud represents an infall per unit area of about the theoretically expected rate, and ~0.1–0.2 times the amount required for the whole Galaxy.

Our result is based on spectra of the Seyfert galaxy Markarian 290, which samples the high-velocity cloud (HVC), complex C (ref. 9). High-velocity clouds consist of neutral hydrogen moving at velocities incompatible with a simple model of differential galactic rotation⁹. Complex C is approaching the Sun at ~150 km s⁻¹. Assuming the tangential velocity is not extremely large, complex C appears to be falling towards the Galactic plane at 50–100 km s⁻¹, after correcting for differential galactic rotation. Figure 1 presents an all-sky map of the HVCs, with complex C highlighted. The spectra are collected in Fig. 2.

S II absorption associated with complex C was observed using the Goddard High Resolution Spectrograph on the Hubble Space Telescope. S⁺ provides a good measure of the intrinsic metallicity of a hydrogen cloud, as sulphur is one of a few elements in the interstellar medium not depleted onto dust, while S⁺ is the domi-

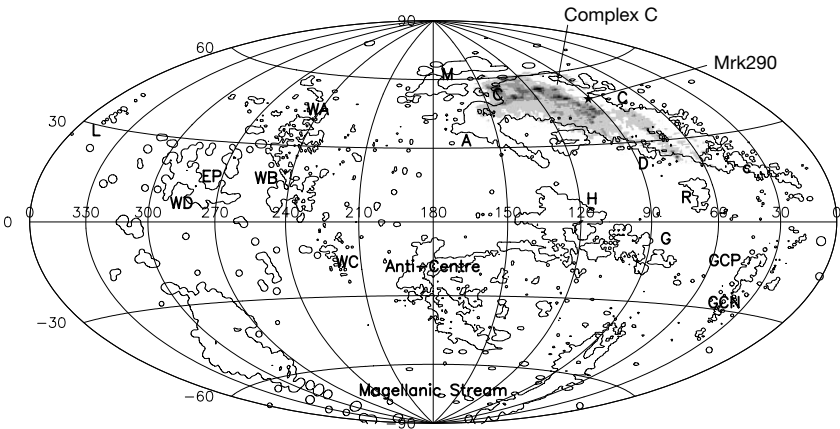


Figure 1 All-sky map of the high-velocity clouds (HVCs) in Aitoff projection, from 21-cm surveys⁹. HVCs consist of gas moving at velocities incompatible with a simple model of differential galactic rotation. The Galactic equator runs through the middle, the Galactic anti-centre is in the centre of the plot. The contour level is at a hydrogen column density of

$2 \times 10^{18} \text{ cm}^{-2}$. The names of the main complexes (from ref. 9) are indicated. HVC complex C is highlighted; the darkest pixels correspond to a column density of $1 \times 10^{20} \text{ cm}^{-2}$. The position of Markarian 290 is indicated (Mrk290).

Table 1 Systematic errors in the determination of the sulphur abundance

Effect	Relevant observational quantity	Observed value	Possible range	Difference in derived abundance
Geometry	See below			-0.012 to +0.000
Temperature	$I(\text{S II})/I(\text{H}\alpha)$	0.03	0.10–0.01	-0.003 to +0.002
Saturation	FWHM (km s ⁻¹)	15	14–10	-0.000 to +0.009
Ionization	$N(\text{S}^{2+})/N(\text{S}^+)$ in H ⁺ region	0.5	0.0–0.7	-0.006 to +0.003
H α fine structure	$I(\text{H}\alpha)$ (Rayleigh)	0.187	0.094–0.280	-0.003 to +0.004
Component structure	$N(\text{H I})$ (10 ¹⁸ cm ⁻²)	92	92–53	-0.000 to +0.060
Distance	D (kpc)	10	25–5	-0.009 to +0.004
Add linearly				-0.033 to +0.082
Add in quadrature (=3 σ)				-0.017 to +0.061

The geometry is usually assumed to be a neutral core and ionized envelope with constant density. Alternatives include a gaussian density profile in the line-of-sight and/or constant partial ionization. A full-width at half-maximum (FWHM) of 10 km s⁻¹ is the minimum line width consistent with the equivalent width errors. S²⁺ only occurs where hydrogen is ionized; in diffuse ionized gas the ratio $N(\text{S}^{2+})/N(\text{S})$ is typically <0.5 (ref. 27). Si λ 1262.86 is not detected, yielding a lower limit $N(\text{S}^{2+})/N(\text{S, tot}) > 0.4$; however, S⁺ has low abundance in the diffuse ISM¹⁰, as it has a low ionization potential (10.4 eV). The possible range in H α emission allows for structure on small angular scales. The largest uncertainty is associated with the component structure, as the S II spectrum does not show separate absorptions at the velocities of the two H I components. If the S II absorption were only associated with the ~138 km s⁻¹ H I component (and if we associate half of the H α emission with it), the implied sulphur abundance would be 0.15 times solar. The final systematic error is calculated by summing in quadrature the maximum changes associated with the listed effects. Even if all systematic effects were to work in the same direction, the sulphur abundance must still be in the range 0.06–0.17 times solar to be compatible with the measurements.

nant ionization stage in neutral and weakly photoionized gas¹⁰, such as exists in complex C. It is assumed that the heavy elements occur in the same proportions in the HVC as in stars. Previously, only depleted elements had been observed in HVCs⁹, except in one cloud which is associated with the Magellanic stream¹¹ and not currently falling in.

A good abundance determination requires as a reference the total (neutral plus ionized) hydrogen column density. As the S II absorption takes place in a pencil-beam path through the HVC, the hydrogen should be measured at the highest-possible angular resolution. The neutral component was found by combining data from the Westerbork Synthesis Radio Telescope (WSRT; 2' beam) and the Effelsberg telescope (9' beam). Figure 2d shows the combined spectrum, Fig. 2e shows the Effelsberg spectrum separately. H α emission from ionized hydrogen was observed using the Wisconsin H α Mapper (WHAM¹²; Fig. 2g), although only at 1° resolution. To determine the average ionized fraction, an H I spectrum at 1° resolution was created by averaging four spectra from the Leiden-Dwingeloo survey¹³ (Fig. 2f). We also used WHAM to set a limit to the [S II] 6,716 Å emission line (Fig. 2h).

We derive physical parameters by assuming a distance and a geometry, that is, a density and ionization structure in the sightline. Using the intensity ratio $I([S II])/I(H\alpha)$ to estimate the temperature then allows us to convert the H α emission measure to an H⁺ column density¹⁴. We find a sulphur abundance $A(S) = 0.089 \pm 0.024^{+0.020}_{-0.005}$ times the solar value¹⁵, a total hydrogen density $n = 0.08 \pm 0.01 (D/10 \text{ kpc})^{-0.5} \text{ cm}^{-3}$, an ionization fraction $N(H^+)/N(H, \text{tot}) = 17 \pm 5^{+7}_{-1} \%$, a temperature $T = 7,300 \pm 900^{+1,500}_{-1,000} \text{ K}$, and a thermal pressure $P = 580 \pm 130 \pm 130 (D/10 \text{ kpc})^{-0.5} \text{ K cm}^{-3}$. Here, D is the distance to the cloud. The first 1 σ error comes from the statistical error in the observations. The second is the 1 σ systematic error, as summarized in Table 1.

To show that complex C is associated with the Milky Way requires an estimate of its distance. We have derived a lower limit¹⁶ of 5 kpc (that is, >3 kpc above the Galactic disk). A definite upper limit is not known, but plausible arguments can be made that complex C is relatively nearby. First, it has large angular size (20° × 90°). Second, if the H⁺ is produced by high-energy photons escaping the disk, then H α brightness is a measure of height above the Galactic plane. Complex C has an H α brightness comparable to that of another HVC which has a distance of 4–10 kpc (ref. 17). Third, we can assume pressure equilibrium between hot gas in the Galactic halo and the HVC, deriving the external pressure from the observed temperature and density of the X-ray halo¹⁸, combined with a semi-empirical halo model¹⁹. This yields 7 kpc, but up to 25 kpc is compatible with 1 σ errors on the parameters. We note that for a velocity of 50–100 km s⁻¹ relative to the surrounding halo gas (below the sound speed of ~120 km s⁻¹ at $T = 10^6 \text{ K}$), the ram pressure is a factor of 7–2 less than the thermal pressure.

We derive a mass of $6.2 \times 10^6 (D/10 \text{ kpc})^2 M_\odot$ for complex C (where $M_\odot$ is the solar mass), assuming no molecular hydrogen²⁰, constant $N(H^+)/N(H, \text{tot})$, and a 28% mass fraction of helium. Using the possible range of its vertical velocity, and dividing the mass by the time to impact gives an associated mass flow of 0.002–0.004 $(D/10 \text{ kpc})^{-1} M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$, or 0.08–0.19 $(D/10 \text{ kpc}) M_\odot \text{ yr}^{-1}$. To solve the G-dwarf problem, models require an inflow of gas with one-tenth solar metallicity at a current rate² of 0.004 $M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$, similar to the rate associated with complex C. To get the required total infall of ~1 $M_\odot \text{ yr}^{-1}$, there should be 5–10 $(D/10 \text{ kpc})^{-1}$ similar objects around the Milky Way; several other HVCs are obvious candidates (see Fig. 1).

Given the position, velocity and metallicity of complex C, what can we say about its origin? It is not possible that it originated in the Galactic disk at galactocentric distances less than about 15 kpc (the Sun is at 8 kpc), as part of a Galactic fountain²¹. In that case, near-solar metallicity is expected. It is also unlikely that it originated in the outer disk. The lowest observed sulphur abundance there is 0.11

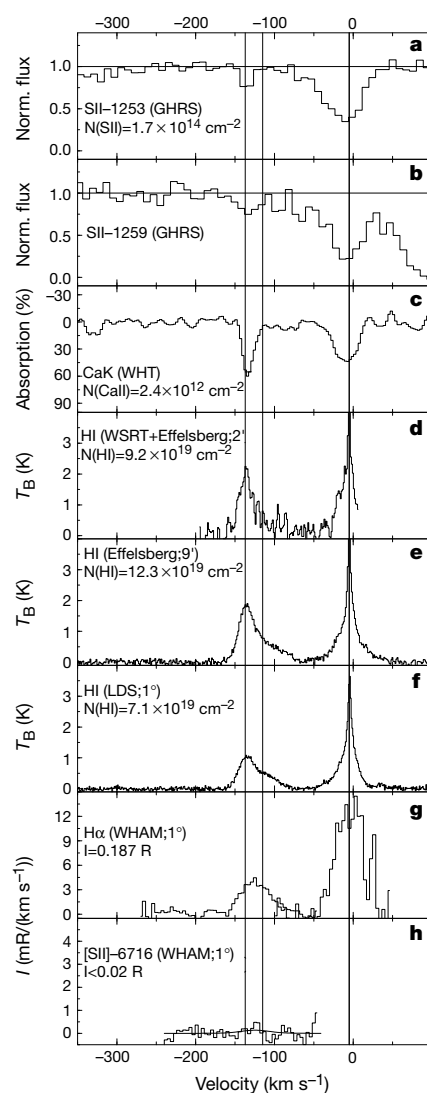


Figure 2 Spectra on or centred on Markarian 290, aligned in velocity. Vertical lines are drawn at velocities (km s⁻¹) of -138, -115 and 0 relative to the local standard of rest (LSR). The first two components belong to the HVC, complex C, whereas the low-velocity component corresponds to gas near the Sun; the HVC approaches about 100 km s⁻¹ faster than can easily be understood from differential galactic rotation. Panels **a** and **b** show two of the three S II absorption lines, observed with the Goddard High-Resolution Spectrograph (GHRS), after normalization to the continuum (15 km s⁻¹ velocity resolution; the 1,250 Å line is confused with absorption intrinsic to Markarian 290). Equivalent widths were converted to column density using the best f -values, and assuming an intrinsic full-width at half-maximum (FWHM) of 15 km s⁻¹ based on the resolved Ca II lines. The values for the 1,253 and 1,259 Å lines were weighted by the signal-to-noise ratio in the continuum to derive $N(S^+) = (1.68 \pm 0.35) \times 10^{14} \text{ cm}^{-2}$. **c**, Ca II spectrum observed with the William Herschel Telescope (WHT), from ref. 13, at 6 km s⁻¹ velocity resolution. **d–f**, H I spectrum observed with different beams, at 1 km s⁻¹ velocity resolution. The 2' spectrum (**d**) yields $N(H I) = (91.7 \pm 7.2) \times 10^{18} \text{ cm}^{-2}$ toward Mrk290, the 9' spectrum (**e**) shows the component structure, while the 1° spectrum (**f**, $N(H I) = (70.6 \pm 4.0) \times 10^{18} \text{ cm}^{-2}$, from the Leiden-Dwingeloo Survey, LDS) is compared to the H α data to derive the ionization fraction. **g**, H α spectrum (12 km s⁻¹ velocity resolution, observed with the Wisconsin H α Mapper, WHAM), showing an HVC intensity of $0.187 \pm 0.01 \text{ R}$ (1 rayleigh is $10^6/4\pi \text{ photons cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$). This combines separate 20-min ($v < -50 \text{ km s}^{-1}$) and 30-s ($v > -100 \text{ km s}^{-1}$) integrations; the feature at about +20 km s⁻¹ is fixed-pattern noise. **h**, [S II] 6,716 of spectrum (12 km s⁻¹ velocity resolution), in which the HVC is not detected, overlaid with a curve for which $I([S II])/I(H\alpha) = 0.03$. We derive a limit $I([S II])/I(H\alpha) < 0.1$ (3 σ). For Galactic disk gas with solar abundances, this ratio is typically ~0.3, with a minimum of ~0.2 (ref. 27). Since for complex C the ratio $N(S^+)/N(H^+)$ is ~0.1 solar, an emission ratio ~0.03 is expected, consistent with our non-detection.

times solar⁵, but to move $\sim 10^7 M_{\odot}$ of gas inward and to a position >3 kpc above the disk requires more energy than seems available.

A number of authors (see ref. 9 for an overview), have proposed that the Local Group (a cluster of galaxies, of which our Galaxy is a member) is filled with hydrogen clouds with typical hydrogen masses of $10^7 M_{\odot}$. Adding large amounts of dark matter (10 times as much mass as in hydrogen) could make such clouds stable²². In the early universe, these mini-galaxies would pick up heavy elements from the then-nearby large galaxies as the latter vigorously formed stars. Complex C could be such an object, whose orbit brought it close to the Milky Way at the present time. A problem with this model is that similar clouds were not detected in other galaxy groups during a sensitive survey done at Arecibo²³.

A more probable origin was proposed by Oort²⁴, who argued that the Milky Way is still forming. In this model, much gas was left over after the Milky Way's original formation, and this gas slowly accretes over time. Complex C (and a few other HVCs) would be among the few still-to-be-accreted objects. Its heavy-element content is then understood as a contamination created after it started interacting with gas in the Galactic halo.

Alternatively, HVCs like complex C may consist of gas tidally stripped from nearby dwarf galaxies when the latter's orbit brought them near the Milky Way. The heavy elements would then have been formed in the dwarf. These dwarfs could subsequently have evolved into the present dwarf spheroidal galaxies²⁵, or their stars might have merged with those of the Milky Way. That such processes still operate is shown by the presence of the Magellanic stream²⁶, a tidal tail torn out of the Small Magellanic Cloud 2 Gyr ago. Combining the theoretical understanding of the chemical evolution and formation of the Milky Way with the constraints set by observations of other galaxy groups, the current content of the Local Group as well as the position, velocity and metallicity of complex C, we conclude that rather than having been assembled in the early universe, it is more probable that the formation of the Milky Way is still continuing. This is fed by gas either left over from the original formation of the Milky Way or stripped from Local Group dwarf galaxies. □

Received 14 June; accepted 4 October 1999.

1. Van den Bergh, S. The frequency of stars with different metal abundances. *Astron. J.* **67**, 486–490 (1962).
2. Giovagnoli, A. & Tosi, M. Chemical evolution models with a new stellar nucleosynthesis. *Mon. Not. R. Astron. Soc.* **273**, 499–504 (1995).
3. Pagel, B. E. J. *Nucleosynthesis and Chemical Evolution of Galaxies* (Cambridge Univ. Press, 1997).
4. Skillman, E. D., Bomans, D. J. & Kobulnicky, H. A. Interstellar medium abundances in the Pegasus dwarf irregular galaxy. *Astrophys. J.* **474**, 205–216 (1997).
5. Rudolph, A. L., Simpson, J. P., Haas, M. R., Erickson, E. F. & Fich, M. Far-infrared abundance measurements in the outer galaxy. *Astrophys. J.* **489**, 94–101 (1997).
6. Ferguson, A. M. N., Gallagher, J. S. & Wyse, R. F. G. The extreme outer regions of disk galaxies. I. Chemical abundances of HII regions. *Astron. J.* **116**, 673–690 (1998).
7. Lu, L., Sargent, W. L. W. & Barlow, T. A. The N/Si abundance ratio in 15 damped Ly α galaxies: implications for the origin of nitrogen. *Astron. J.* **115**, 55–61 (1998).
8. Pettini, M., Ellison, S. L., Steidel, C. C. & Bowen, D. V. Metal abundances at $z < 1.5$: fresh clues to the chemical enrichment history of damped Ly α systems. *Astrophys. J.* **510**, 576–589 (1999).
9. Wakker, B. P. & van Woerden, H. High-velocity clouds. *Annu. Rev. Astron. Astrophys.* **35**, 217–266 (1997).
10. Savage, B. D. & Sembach, K. R. Interstellar abundances from absorption-line observations with the Hubble Space Telescope. *Annu. Rev. Astron. Astrophys.* **34**, 279–329 (1996).
11. Lu, L. *et al.* The metallicity and dust content of HVC 287.5 + 22.5 + 240: evidence for a Magellanic Clouds origin. *Astron. J.* **115**, 162–167 (1998).
12. Reynolds, R. J., Tuft, S. L., Haffner, L. M., Jaehnig, K. & Percival, J. W. The Wisconsin H-alpha Mapper (WHAM): A brief review of performance characteristics and early scientific results. *Publ. Astron. Soc. Aust.* **15**, 14–18 (1998).
13. Hartmann, D. & Burton, W. B. *Atlas of Galactic Neutral Hydrogen* (Cambridge Univ. Press, 1997).
14. Wakker, B. P. *et al.* in *Strömlo Workshop on High-velocity Clouds* (eds Gibson, B. K. & Putman, M. E.) 26–37 (ASP Conf. Ser. 166, Astronomical Soc. of the Pacific, San Francisco, 1999).
15. Anders, E. & Grevesse, N. Abundances of the elements—Meteoritic and solar. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).
16. van Woerden, H., Peletier, R. F., Schwarz, U. J., Wakker, B. P. & Kalberla, P. M. W. in *Strömlo Workshop on High-velocity Clouds* (eds Gibson, B. K. & Putman, M. E.) 1–25 (ASP Conf. Ser. 166, Astronomical Soc. of the Pacific, San Francisco, 1999).
17. van Woerden, H., Schwarz, U. J., Peletier, R. F., Wakker, B. P. & Kalberla, P. M. W. A confirmed location in the Galactic halo for the high-velocity cloud 'Chain A'. *Nature* **400**, 138–141 (1999).
18. Snowden, S. L., Egger, R., Finkbeiner, D. P., Freyberg, M. J. & Plucinsky, P. P. Progress on establishing

the spatial distribution of material responsible for the 1/4 keV soft x-ray diffuse background local and halo components. *Astrophys. J.* **493**, 715–729 (1998).

19. Wolfire, M. G., McKee, C. F., Hollenbach, D. & Tielens, A. G. G. M. The multiphase structure of the galactic halo: high-velocity clouds in a hot corona. *Astrophys. J.* **453**, 673–684 (1995).
20. Wakker, B. P., Murphy, E. M., van Woerden, H. & Dame, T. M. A sensitive search for molecular gas in high-velocity clouds. *Astrophys. J.* **488**, 216–223 (1997).
21. Bregman, J. N. The galactic fountain of high-velocity clouds. *Astrophys. J.* **236**, 577–591 (1980).
22. Blitz, L., Spergel, D., Teuben, P., Hartmann, D. & Burton, W. B. High velocity clouds: building blocks of the Local Group. *Astrophys. J.* **514**, 818–843 (1999).
23. Zwaan, M. A., Briggs, F. H., Sprayberry, D. & Sorar, E. The HI mass function of galaxies from a deep survey in the 21-cm line. *Astrophys. J.* **490**, 173–186 (1997).
24. Oort, J. H. The formation of galaxies and the origin of the high velocity hydrogen. *Astron. Astrophys.* **7**, 381–404 (1970).
25. Yanny, B. & York, D. G. Emission-line objects near quasi-stellar object absorbers. III Clustering and colors of moderate redshift HII regions. *Astrophys. J.* **391**, 569–576 (1992).
26. Gardiner, L. T. & Noguchi, M. N-body simulations of the Small Magellanic Cloud and the Magellanic Stream. *Mon. Not. R. Astron. Soc.* **278**, 191–208 (1996).
27. Haffner, L. M., Reynolds, R. J. & Tuft, S. L. WHAM observations of H α , SII, and NII toward the Perseus Arm: probing the physical conditions of the WIM. *Astrophys. J.* **523**, 223–233 (1999).
28. Verner, D. A., Barthel, P. D. & Tytler, D. Atomic data for absorption lines from the ground level at wavelengths greater than 228 Å. *Astron. Astrophys.* **108**, 287–340 (1994).
29. Wakker, B. P., van Woerden, H., Schwarz, U. J., Peletier, R. F. & Douglas, N. G. The Ca⁺ abundance of HVC complex C. *Astron. Astrophys.* **306**, L25–L28 (1996).

Acknowledgements

This work was supported by the Space Telescope Science Institute. The Hubble Space Telescope is operated by the Association of Universities for Research in Astronomy, Inc. The Effelsberg Telescope belongs to the Max Planck Institute for Radio Astronomy in Bonn. The Westerbork Radio Observatory is operated by the Netherlands Foundation for Research in Astronomy (ASTRON/NFRA) with financial support from NWO.

Correspondence and requests for materials should be addressed to B.P.W. (e-mail: wakker@astro.wisc.edu).

Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations

Daniel Gottesman^{*†} & Isaac L. Chuang[‡]

^{*} Theoretical Astrophysics T-6, MS B-288, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

[†] Microsoft Research, One Microsoft Way, Redmond, Washington 98055, USA

[‡] IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120, USA

Algorithms such as quantum factoring¹ and quantum search² illustrate the great theoretical promise of quantum computers; but the practical implementation of such devices will require careful consideration of the minimum resource requirements, together with the development of procedures to overcome inevitable residual imperfections in physical systems^{3–5}. Many designs have been proposed, but none allow a large quantum computer to be built in the near future⁶. Moreover, the known protocols for constructing reliable quantum computers from unreliable components can be complicated, often requiring many operations to produce a desired transformation^{3–5,7,8}. Here we show how a single technique—a generalization of quantum teleportation⁹—reduces resource requirements for quantum computers and unifies known protocols for fault-tolerant quantum computation. We show that single quantum bit (qubit) operations, Bell-basis measurements and certain entangled quantum states such as Greenberger–Horne–Zeilinger (GHZ) states¹⁰—all of which are within the reach of current technology—are sufficient to construct a universal quantum computer. We also present systematic constructions

for an infinite class of reliable quantum gates that make the design of fault-tolerant quantum computers much more straightforward and methodical.

Quantum teleportation is a scheme by which the state of a qubit can be transported from one point to another by communicating just two classical bits, provided that the sender and the receiver have previously shared halves of a specific two-qubit entangled state. In detail, this is how it works (Fig. 1): a single-qubit state $|\alpha\rangle = a|0\rangle + b|1\rangle$ is prepared, along with a two-qubit entangled state known as an 'Einstein-Podolsky-Rosen' (EPR) state, $|\Psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. The top two qubits, $|\alpha\rangle$ and one qubit of $|\Psi\rangle$, belong to the sender, who measures them in the Bell basis (see Fig. 1 legend). No matter what the input state $|\alpha\rangle$ is, xy is a uniformly distributed random two-bit classical result¹¹. This measurement leaves the third qubit, belonging to the receiver, in the state $R_{xy}|\alpha\rangle$, where R_{xy} is either the identity (no transformation) if $xy = 00$, or a single qubit Pauli operation $R_{10} = X$ (a bit flip $|a\rangle \rightarrow |a \oplus 1\rangle$), $R_{01} = Z$ (a phase flip $|a\rangle \rightarrow (-1)^a|a\rangle$), or $R_{11} = Y$ (bit + phase flip $|a\rangle \rightarrow i(-1)^a|a \oplus 1\rangle$)¹². To complete the teleportation the sender transmits the two classical bits xy to the receiver, who uses them to apply a correction procedure $R_{xy}^\dagger$, which is just the inverse of R_{xy} , thus reconstructing an output $|\alpha\rangle$ identical to the original input.

We now consider several features of this procedure which help the intuitive understanding of the result we report here. Teleportation utilizes a specific entangled state $|00\rangle + |11\rangle$ as a resource in accomplishing its aims; what happens when this state is faulty? In general, the teleportation procedure would fail, but sometimes it would fail in a useful manner. If $|\Psi\rangle$ is replaced by $U|\Psi\rangle$, where U is some non-trivial quantum operation, the teleportation procedure produces an output which is not identical to the input. We show below that for certain U , the procedure can be modified so that it produces precisely $U|\alpha\rangle$, that is, the output is a transformed version of the input, where the transformation is determined by the pre-computed entangled state used by the procedure. We shall describe this act as teleporting a state 'through' U . The important points are that (1) the transformation U is potentially useful—it could be an otherwise difficult-to-implement logic gate, (2) the modified teleportation procedure is simple—it does not require complicated gates, and (3) the resource $U|\Psi\rangle$ can be reliably constructed, even using unreliable gates.

An example illustrates how this works. The Hadamard gate $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is an important elementary operation for quantum computers, but it cannot be constructed from the Pauli gates X , Y and Z . It turns out that H can be accomplished using teleportation. Instead of $|\Psi\rangle$, use $(I \otimes H)|\Psi\rangle$ as the input entangled state; $I \otimes H$ means 'do nothing to the first qubit, and apply H to the second qubit'. This second qubit is the one possessed by the receiver. The

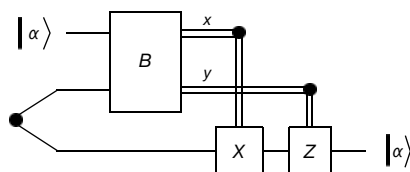


Figure 1 Quantum circuit for teleportation. Time proceeds from left to right, and $\langle$ denotes the EPR state $|\Psi\rangle$. The double wires carry classical bits, and the single wires, qubits. The box B represents measurement in the Bell basis; that is, if the two qubits entering B are found to be $|00\rangle + |11\rangle$ (leaving out the $1/\sqrt{2}$ normalization for clarity), then the outputs $xy = 00$; for $|01\rangle + |10\rangle$, $xy = 10$; for $|00\rangle - |11\rangle$, $xy = 01$; and for $|01\rangle - |10\rangle$, $xy = 11$; one of the four possibilities is guaranteed to be true because the states form a complete basis for all possible two-qubit states. Measurement, in its guise as an interface between the quantum and classical worlds, is generally considered to be an irreversible operation, destroying quantum information and replacing it with classical information. But this circuit demonstrates that in certain carefully designed cases this need not be true, as teleportation uses measurement to transfer states from one place to another.

first step of the usual teleportation procedure leaves the receiver with the qubit $HR_{xy}|\alpha\rangle$. We note that H appears to have been performed after R_{xy} even though in practice the Bell-basis measurement is done long after the preparation of $(I \otimes H)|\Psi\rangle$; this time-reversal is a property of using entangled states^{5,13}. This seems to be rather inconvenient, because what we really want is $H|\alpha\rangle$. But it happens that $HR_{xy} = R_{x'y'}H$; the Hadamard gate, when commuted through a Pauli gate, produces a Pauli gate. It may be a different gate, but no matter—we know *a priori* what the classical mapping between xy and $x'y'$ is (in fact, in this case, $x'y' = xy$), and thus when the receiver applies $R_{x'y'}^\dagger$, the output state $H|\alpha\rangle$ is obtained, as desired.

A more useful and important gate is the controlled-NOT (CNOT) gate, which acts on two qubits, a control and a target qubit, and flips the target whenever the control qubit is a $|1\rangle$ ($|a, b\rangle \rightarrow |a, a \oplus b\rangle$). We can teleport a state through a CNOT using the quantum circuit shown in Fig. 2. This can be verified by direct computation, but it is easier to understand by realizing that $|\chi\rangle$ can be created by simply performing a CNOT on two EPR pairs (Fig. 3). Combining this circuit with the circuit in Fig. 2, we can see what is happening. Two input qubits are teleported through the CNOT; the correction now depends on four classical bits instead of two. This construction enables CNOT gates to be performed between two qubits, using only classically controlled single-qubit operations, prior entanglement, and Bell-basis measurements. As any quantum computation can be performed using single-qubit operations and CNOT gates¹⁴, this demonstrates an alternative universal set of operations for quantum computation, which does not require any two-qubit operation except the measurement. Moreover, the entangled state $|\chi\rangle$ that is used as a resource can be readily created from two pairs of GHZ¹⁰ states (Fig. 3).

Generalizing this procedure allows us to construct an entire hierarchy of gates through which states can be teleported, and in a fault-tolerant fashion. It is no accident that H and CNOT transform Pauli gates into Pauli gates; the set of gates which does so is known as the Clifford group. It is important in the theory of quantum error-correcting codes and fault-tolerance^{5,15}. We now define precisely what operations U can be constructed via teleportation, and how this is done reliably.

Consider an n -qubit state $|\psi\rangle$, in which each qubit is encoded using a stabilizer code, such as the 7-qubit CSS code^{16,17}. 0 and 1 shall now represent the corresponding encoded qubit states. Let $|\Psi^n\rangle$ be the $2n$ -(encoded) qubit Bell state $(|00\rangle + |11\rangle)^{\otimes n}$ (normalizations suppressed for clarity), rearranged so that the first n labels (the

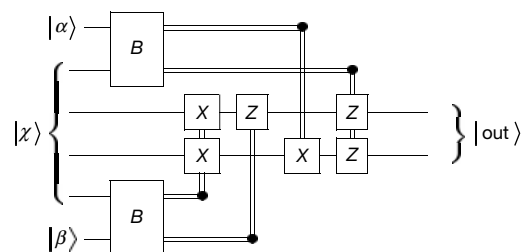


Figure 2 Quantum circuit for teleporting two qubits through a controlled-NOT gate. The output is $|\text{out}\rangle = \text{CNOT}|\beta\rangle\alpha\rangle$, where the inputs $|\alpha\rangle = a|0\rangle + b|1\rangle$ and $|\beta\rangle = c|0\rangle + d|1\rangle$ are two arbitrary single qubit states. The CNOT gate has $|\beta\rangle$ as its control, and $|\alpha\rangle$ as its target. The special entangled state $|\chi\rangle$ is $(|00\rangle + |11\rangle)(|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$. We recall that multiple qubits can be teleported simply by replicating instances of the single-qubit procedure. Normally, for two qubits, the receiver would obtain the state $R_{x_1y_1}R_{x_2y_2}|\beta, \alpha\rangle$, but because we have replaced the usual EPR pairs by $|\chi\rangle$, the receiver instead obtains $\text{CNOT}R_{x_1y_1}R_{x_2y_2}|\beta, \alpha\rangle$. Now, just as with the Hadamard gate, CNOT has the fortunate property that $\text{CNOT}R_{x_1y_1}R_{x_2y_2} = R_{x'_1y'_1}R_{x'_2y'_2}\text{CNOT}$. This means that the receiver can again apply a modified correction procedure (shown here) to obtain $\text{CNOT}|\beta, \alpha\rangle$.

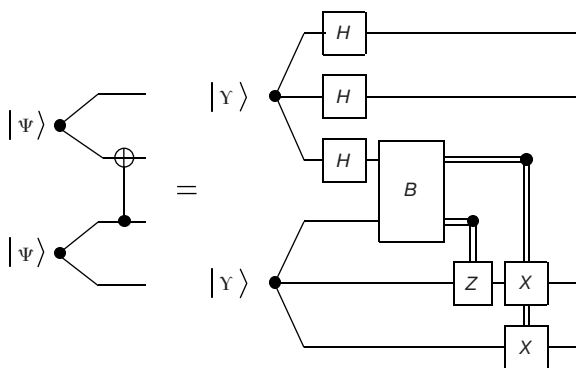


Figure 3 Quantum circuit to create the $|\chi\rangle$ state. It uses two EPR pairs (left) or two GHZ states $|\Upsilon\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$ (right). H is the Hadamard gate.

upper qubits) represent qubits from half of each EPR pair, and the last n (the lower qubits) the other halves. In other words, $(I \otimes U)|\Psi^n\rangle$ (where I is the identity on n qubits) is U acting on the lower qubits of all the EPR pairs.

The goal of fault-tolerant computation is to perform logical operations using quantum gates while restricting the propagation of errors among the physical qubits, which can compromise the code's ability to correct errors. The usual method for doing this is to only perform transversal gates on the code—that is, gates which interact qubits in one code block only with corresponding qubits in other code blocks. While errors may then propagate between blocks, they cannot propagate within blocks, so a single faulty gate can only cause a single error in any given block of the code.

Operators from the Pauli group (such as X , Y and Z) can readily be performed on logical qubits which are encoded with a stabilizer code¹⁸. Let C_1 represent the Pauli group. C_2 , the Clifford group, will be the set of gates which map Pauli operators into Pauli operators under conjugation. Through an appropriate sequence of gates and measurements, any C_2 operation can also be performed on any stabilizer code¹⁸.

More difficult to perform are gates in the class defined as $C_3 \equiv \{U|UC_1U^\dagger \subseteq C_2\}$. C_3 contains gates such as the Toffoli gate (controlled-controlled-NOT), the $\pi/8$ gate (rotation about the Z -axis by an angle $\pi/4$), and the phase gate ($\text{diag}(1, 1, i)$). Fault-tolerant constructions of these gates are known^{7,8,19}, but they are *ad hoc* and do not generalize easily.

However, our teleportation construction provides a straightforward way to produce any gate in C_3 , as shown in Fig. 4. For $U \in C_3$, first construct the state $|\Psi_U^n\rangle = (I \otimes U)|\Psi^n\rangle$. Next, take the input state $|\Psi\rangle$ and do Bell basis measurements on this and the n upper qubits of $|\Psi_U^n\rangle$, leaving n qubits in the state $|\psi_{\text{out}}\rangle = UR_{xy}|\psi\rangle = R'_{xy}U|\psi\rangle$, where R_{xy} is an operator in C_1 which depends on the (random) Bell-basis measurement outcomes xy , and $R'_{xy} = UR_{xy}U^\dagger$ is an operator in C_2 . As R'_{xy} is in the Clifford group, it can be performed—or more precisely, inverted—fault-tolerantly.

Creating $|\Psi_U^n\rangle$ may appear, at first sight, to be as difficult as performing U . However, constructing specific known states is easier than doing operations on unknown states: if the construction of an ancilla fails, little is lost by discarding the ruined state and starting again. This option is not available when we try to perform logical gate operations on actual data.

Our construction of quantum gates using teleportation offers possibilities for relaxing experimental constraints on realizing quantum computers. For example, using single photons as qubits and current optical technology, one can perform nearly perfect Bell-basis measurements²⁰, quantum teleportation²¹, almost create GHZ states²², and certainly perform single-qubit operations²³. Thus, given GHZ states, quantum computers might be constructed almost completely from linear optical components. Our results

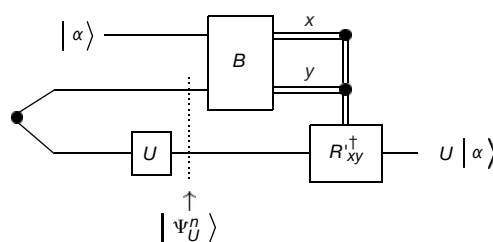


Figure 4 Quantum circuit to perform U in a fault-tolerant manner using quantum teleportation. In general, this works for any $U \in C_k$, defined as $C_k = \{U|UC_1U^\dagger \subseteq C_{k-1}\}$. Then $R'_{xy} \in C_{k-1}$. $|\Psi_U^n\rangle$ must also be prepared fault-tolerantly. To do this, we note that the state $|\Psi_U^n\rangle$ is the $+1$ eigenvector of the $2n$ operators $X_i \otimes X_i$ and $Z_i \otimes Z_i$ (where X_i and Z_i are X and Z , respectively, acting on the i th upper or lower qubit). A computation shows that $|\Psi_U^n\rangle$ is therefore the $+1$ eigenvector of the operators $M_i = X_i \otimes UX_iU^\dagger$ and $N_i = Z_i \otimes UZ_iU^\dagger$. Furthermore, the eigenvalues of these $2n$ operators completely determine the state, so if all these operators have eigenvalue $+1$, the state actually is the desired one. Therefore, to produce $|\Psi_U^n\rangle$, measure the operators M_i and N_i , which are in the Clifford group C_2 . If the eigenvalue of operator M_i is -1 , perform the Pauli operation $Z_i \otimes I$; if the operator N_i has eigenvalue -1 , perform $X_i \otimes I$. These operators anticommute with M_i and N_i , so move the state from the -1 eigenspace into the $+1$ eigenspace, resulting in the state $|\Psi_U^n\rangle$. The hard part of the preparation is measuring M_i and N_i fault-tolerantly, but this can be done by adapting a trick of Shor's⁷. Although the precise set of gates which form C_k is still under investigation, it is known that every C_k contains interesting gates not in an earlier C_k , such as the $\pi/2^k$ rotations, which appear in Shor's factoring algorithm¹. The union of the C_k s is therefore infinite, even for a fixed number of qubits.

have similar implications for other physical systems, particularly if entangled states can readily be prepared and stored.

The Toffoli gate or another gate in C_3 is required for a universal fault-tolerant quantum computer, and our construction using teleportation offers a conceptual simplification over previous constructions. As with earlier realizations, our construction relies on the ability to create certain ancilla states $|\Psi_U^n\rangle$ which are independent of the data being acted upon. This means they may be prepared off-line, so $|\Psi_U^n\rangle$ are valuable generic quantum resources, perhaps a kind of 'quantum software', which might be considered a commercial commodity that could be manufactured. Even if states $|\Psi_U^n\rangle$ are not available, the construction can in some cases greatly reduce the number of operations needed to assemble the precise gates called for in an algorithm; this is of benefit in efficiently performing quantum computation with realistically imperfect gates. □

Received 15 July; accepted 11 October 1999.

- Shor, P. in *Proc. 35th Annu. Symp. on Foundations of Computer Science* (ed. Goldwasser, S.) 124–134 (IEEE Computer Society Press, Los Alamitos, 1994).
- Grover, L. K. Quantum computers can search arbitrarily large databases by a single query. *Phys. Rev. Lett.* **79**, 4709–4012 (1997).
- Preskill, J. Reliable quantum computers. *Proc. R. Soc. Lond. A* **454**, 385–410 (1998).
- Steane, A. M. Efficient fault tolerant quantum computing. *Nature* **399**, 124–126 (1999).
- Gottesman, D. Theory of fault-tolerant quantum computation. *Phys. Rev. A* **57**, 127–137 (1998).
- Preskill, J. Quantum computing: pro and con. *Proc. R. Soc. Lond. A* **454**, 469–486 (1998).
- Shor, P. W. in *Proc. 37th Annu. Symp. on Foundations of Computer Science* (IEEE Computer Society Press, Los Alamitos, 1996).
- Knill, E., Laflamme, R. & Zurek, W. Resilient quantum computation. *Science* **279**, 342–345 (1998).
- Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and EPR channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
- Greenberger, D., Horne, M., Shimony, A. & Zeilinger, A. Bell's theorem without inequalities. *Am. J. Phys.* **58**, 1131–1143 (1990).
- Brassard, G. in *PhysComp 96* (eds Toffoli, T., Biafore, M. & Leao, J.) 48–50 (New England Complex Systems Inst., Cambridge, Massachusetts, 1996).
- Gottesman, D. in *Group22: Proc. XXII Int. Colloquium on Group Theoretical Methods in Physics* (eds Corney, S. P., Delbourgo, R. & Jarvis, P. D.) 32–43 (International Press, Cambridge, Massachusetts, 1999).
- Nielsen, M. A. & Chuang, I. L. Programmable quantum gate arrays. *Phys. Rev. Lett.* **79**, 321–324 (1997).
- Barenco, A. *et al.* Elementary gates for quantum computation. *Phys. Rev. A* **52**, 3457–3467 (1995).
- Calderbank, A. R., Rains, E. M., Shor, P. W. & Sloane, N. J. A. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.* **78**, 405–408 (1997).
- Steane, A. M. Multiple particle interference and quantum error correction. *Proc. R. Soc. Lond. A* **452**, 2551–2576 (1996).

17. Steane, A. M. Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77**, 793–797 (1996).
18. Gottesman, D. *Stabilizer Codes and Quantum Error Correction*. Thesis, California Inst. of Technol. (1997).
19. Boykin, P. O., More, T., Pulver, M., Roychowdhury, V. & Vatan, F. On universal and fault-tolerant quantum computing. Preprint quant-ph/9906054 (cited June 1999) at (<http://xxx.lanl.gov>) (1999).
20. Kwiat, P. G. & Weinfurter, H. Embedded Bell-state analysis. *Phys. Rev. A* **58**, R2623–R2626 (1998).
21. Bouwmeester, D. *et al.* Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
22. Bouwmeester, D. *et al.* Observation of three-photon Greenberger-Horne-Zeilinger entanglement. *Phys. Rev. Lett.* **82**, 1345–1349 (1999).
23. Chuang, I. L. & Yamamoto, Y. Simple quantum computer. *Phys. Rev. A* **52**, 3489–3496 (1995).

Acknowledgements

We thank C. Bennett for suggesting the concept of “quantum software” to us, and R. Jozsa for pointing out an error in an early version of this manuscript. We also thank J. Kempe, D. Leung, and D. Bacon for helpful discussions. This work was supported in part by DARPA under the NMRQC initiative.

Correspondence and requests for materials should be addressed to I.L.C. (e-mail: ichuang@almaden.ibm.com).

Coupled synthesis and self-assembly of nanoparticles to give structures with controlled organization

Mei Li*, Heimo Schnablegger† & Stephen Mann*

* School of Chemistry, University of Bristol, Bristol BS8 1TS, UK

† Max-Planck-Institut of Colloids and Interfaces, Am Muehlenberg, D-14476 Golm, Germany

Colloidal inorganic nanoparticles have size-dependent optical, optoelectronic and material properties that are expected to lead to superstructures with a range of practical applications^{1,2}. Discrete nanoparticles with controlled chemical composition and size distribution are readily synthesized using reverse micelles and microemulsions as confined reaction media^{3–5}, but their assembly into well-defined superstructures amenable to practical use remains a difficult and demanding task. This usually requires the initial synthesis of spherical nanoparticles, followed by further processing such as solvent evaporation^{6–8}, molecular cross-linking^{9–14} or template-patterning^{15–18}. Here we report that the interfacial activity of reverse micelles and microemulsions can be exploited to couple nanoparticle synthesis and self-assembly over a range of length scales to produce materials with complex organization arising from the interdigitation of surfactant molecules attached to specific nanoparticle crystal faces. We demonstrate this principle by producing three different barium chromate nanostructures—linear chains, rectangular superlattices and long filaments—as a function of reactant molar ratio, which in turn is controlled by fusing reverse micelles and microemulsion droplets containing fixed concentrations of barium and chromate ions, respectively. If suitable soluble precursors and amphiphiles with headgroups complementary to the crystal surface of the nanoparticle target are available, it should be possible to extend our approach to the facile production of one-dimensional ‘wires’ and higher-order colloidal architectures made of metals and semiconductors.

Barium bis(2-ethylhexyl)sulphosuccinate ($\text{Ba}(\text{AOT})_2$) reverse micelles were added to sodium chromate (Na_2CrO_4)-containing NaAOT microemulsion droplets, to give final molar ratios of $[\text{Ba}^{2+}] : [\text{CrO}_4^{2-}] \approx 1$ and water content $w = [\text{H}_2\text{O}] : [\text{NaAOT}] = 10$. This produced a yellow precipitate ~3 h after addition of the reactants at 25 °C. Transmission electron microscopy (TEM) images of samples taken directly from the liquid phase of the micro-

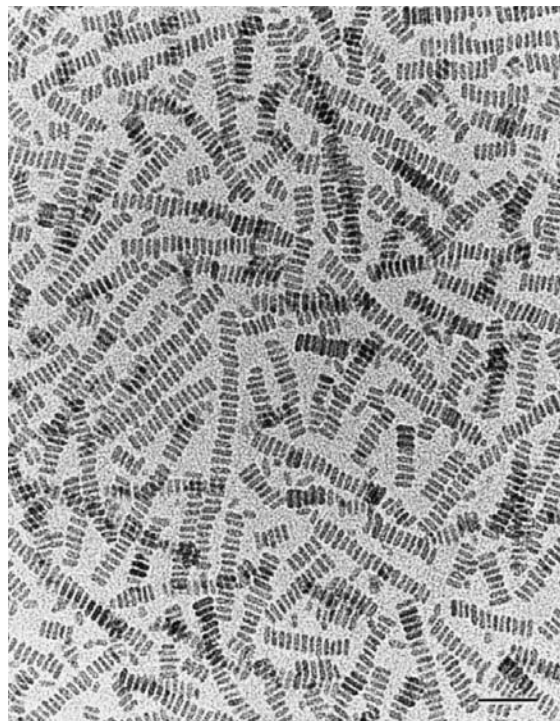


Figure 1 TEM image showing ordered chains of prismatic BaCrO_4 nanoparticles. The nanoparticles were prepared in AOT microemulsions at $[\text{Ba}^{2+}] : [\text{CrO}_4^{2-}]$ molar ratio ≈ 1 and $w = 10$. Scale bar = 50 nm.

emulsion showed chain-like arrays that contained up to 60 nanoparticles (Fig. 1). The colloidal chain structures were 50–500 nm in length and consisted of rectangular-shaped particles that were uniform in length (mean $= 16.0 \pm 1.5$ nm) and width (mean $= 6.0 \pm 0.4$ nm), and preferentially aligned so that the long axis of each particle was perpendicular to the chain direction. Energy dispersive X-ray analysis and powder electron diffraction patterns indicated that the nanoparticles were crystalline BaCrO_4 with an orthorhombic unit cell ($a = 0.911$, $b = 0.554$, $c = 0.734$ nm). Each crystal along the length of the chain was separated by a regular spacing of 2 nm, consistent with the presence of an interdigitated layer of surfactant molecules.

Corresponding TEM studies on the sedimented material showed thin flake-like aggregates of a two-dimensional superlattice constructed from a pseudo-rectangular ($90^\circ \leq \theta \leq 100^\circ$) array of uniformly sized BaCrO_4 nanoparticles that were separated by an interparticle spacing of 2 nm (Fig. 2). Thermal analysis indicated that the superlattices contained ~30% by weight of surfactant. Images of tilted lattices indicated that the particles were prismatic and identical to those present in the chain motif, and aligned with their long axis perpendicular to the plane of the superlattice. Viewed in-plane, the nanoparticles were rectangular in shape with mean dimensions of 6.8 ± 0.6 nm and 5.9 ± 0.5 nm, indicating two different types of side face. Electron diffraction failed to identify unequivocally the crystallographic nature of the side faces because of multiple arcing of the reflections arising from long-range disorder in the air-dried structures. However, patterns recorded from particles in the superlattice structure were indexed according to a superimposition of directions close to the [100] zone axis, which indicated that the prismatic crystals were single-domain particles elongated along the crystallographic a axis. This was consistent with a morphology based on a set of {100} end faces with {010} and {001} side faces.

Systematic changes in the water content ($5 \leq w \leq 20$), and hence

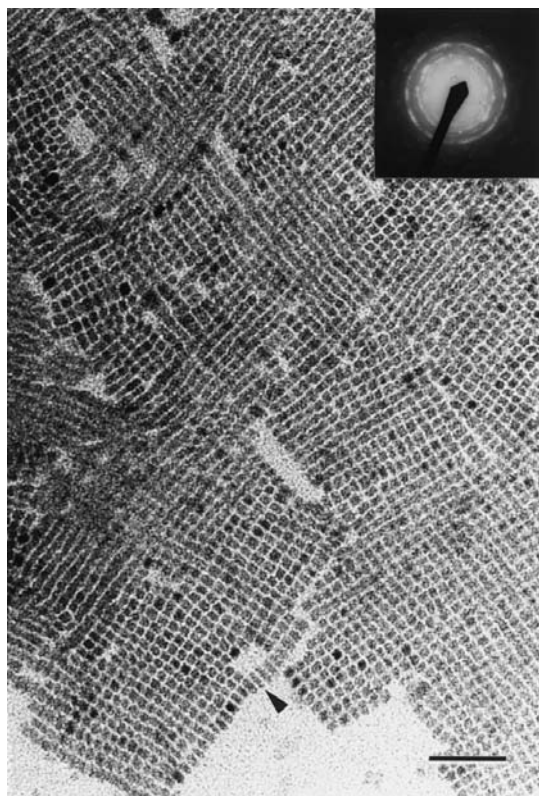


Figure 2 Rectangular superlattice of BaCrO_4 nanoparticles. These were formed by two-dimensional aggregation of nanoparticle chains prepared in AOT microemulsions at $[\text{Ba}^{2+}]:[\text{CrO}_4^{2-}]$ molar ratio ≈ 1 and $w = 10$. Arrow shows dislodged particles revealing the prismatic morphology of individual crystallites. Scale bar = 50 nm. Inset, the electron diffraction pattern gives the superimposition of reflections from zone axes approximately parallel to the $[100]$ direction.

in the diameter of the microemulsion droplets, was used to control the size of the nanoparticles present in both the chain and superlattice structures. For constant values of $[\text{Ba}^{2+}]:[\text{CrO}_4^{2-}] \approx 1$, increases in w gave increased mean particle sizes that were correlated in the chain and superlattice nanostructures (Table 1). The particle width measured for crystals in the individual chains was commensurate with the shorter in-plane dimension of the superlattice particles, suggesting that the larger side faces were preferentially juxtaposed in the stacked linear array.

Dynamic light-scattering (DLS) of unstirred microemulsions, prepared as above at $w = 10$, suggested the presence of microemulsion water droplets with a constant hydrodynamic radius of 5.8 ± 0.5 nm throughout the time course of the experiments. Similar observations were made using small-angle X-ray scattering (SAXS), which detected significant numbers of Ba^{2+} -containing spherical water droplets in the reaction mixtures, even after 25 h at 25 °C. DLS also suggested the appearance of a second component within 3 hours at 25 °C (or 18 h at 20 °C), which progressively increased in scattering intensity over 6 h at 25 °C, until the experiment was terminated due to bulk precipitation. Although the complexity of the system must prevent a conclusive analysis of the DLS data, they do suggest that the second component is associated with the formation of aggregates in the microemulsion fluid. Supporting this interpretation are the results of a TEM analysis of corresponding samples, which show the appearance of the second component in the DLS studies to be concurrent with the appearance of nanoparticle chains. Taken together, the data suggest that spontaneous self-assembly of the BaCrO_4 linear arrays occurs in the microemulsion fluid.

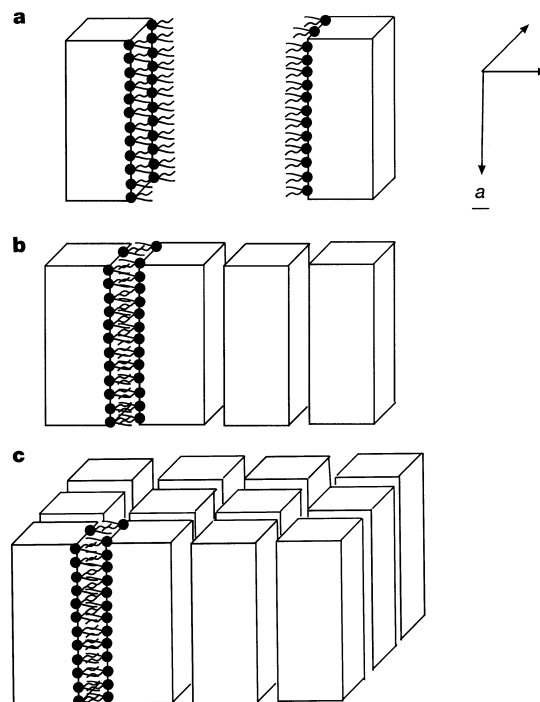


Figure 3 Proposed model for the surfactant-induced self-assembly of nanoparticle chains and superlattices. **a**, Surfactant-coated prismatic BaCrO_4 nanoparticles synthesized by controlled crystallization in microemulsion water droplets. For clarity, only one face is shown with associated surfactant molecules. **b**, Interdigitation of the surfactant monolayers induced as the crystal faces develop in shape and size, resulting in preferential aggregation normal to both the prism long axis (crystallographic a axis) and largest side face. **c**, Aggregation in two dimensions proceeds as the chains develop in length and number.

TEM images of the initial stages of chain assembly showed low-contrast structures with striped patterns, rather than isolated nanoparticles with prismatic morphology, suggesting that BaCrO_4 growth occurs in association and concurrent with the self-organization of stacked micellar aggregates and not through sequential attachment of individual nanoparticles. One possibility is that the linear arrays spontaneously self-assemble in the microemulsion fluid from interactions between the hydrophobic tails of AOT molecules absorbed onto the flat side faces of developing BaCrO_4 prismatic crystallites (Fig. 3). A difference of only $\sim 14 \text{ nm}^2$ in the surface areas of the two sets of side faces appears to be sufficient to stabilize the chain motif before superlattice formation, even though the latter has an increased number of particle–particle connections.

The BaCrO_4 nanoparticle chains and superlattices were formed specifically at molar equivalence ($1:1.4 \leq [\text{Ba}^{2+}]:[\text{CrO}_4^{2-}] \leq 1.4:1$). Under identical conditions, but with an excess of Ba^{2+}

Table 1 Mean dimensions for individual prismatic BaCrO_4 nanoparticles

w value ($\text{H}_2\text{O}/\text{NaAOT}$)	Chain-like stacks (nm)	Superlattices (nm)	Microemulsion radius (nm)
5	$14.5 \pm 1.2 \times 5.6 \pm 0.4$	$6.4 \pm 0.6 \times 5.3 \pm 0.5$	1.5 ± 0.5
10	$16.0 \pm 1.5 \times 6.0 \pm 0.4$	$6.8 \pm 0.6 \times 5.9 \pm 0.5$	2.2 ± 0.6
15	$17.5 \pm 1.4 \times 6.3 \pm 0.4$	$7.4 \pm 0.7 \times 6.1 \pm 0.6$	—
20	$18.3 \pm 1.5 \times 8.4 \pm 0.8$	$10.0 \pm 1.1 \times 8.1 \pm 0.7$	3.2 ± 0.6

The nanoparticles were viewed side-on (chains) and end-on (superlattice) by electron microscopy. The nanostructures were synthesized in AOT/water/isooctane reverse microemulsions at $[\text{Ba}^{2+}]:[\text{CrO}_4^{2-}] \approx 1:1$, $[\text{NaAOT}]:[\text{Ba}(\text{AOT})_2] = 50$, and various water content w values. Sizes of the corresponding microemulsion water droplets were determined by small-angle X-ray scattering. $\sigma < 10\%$ for all data.

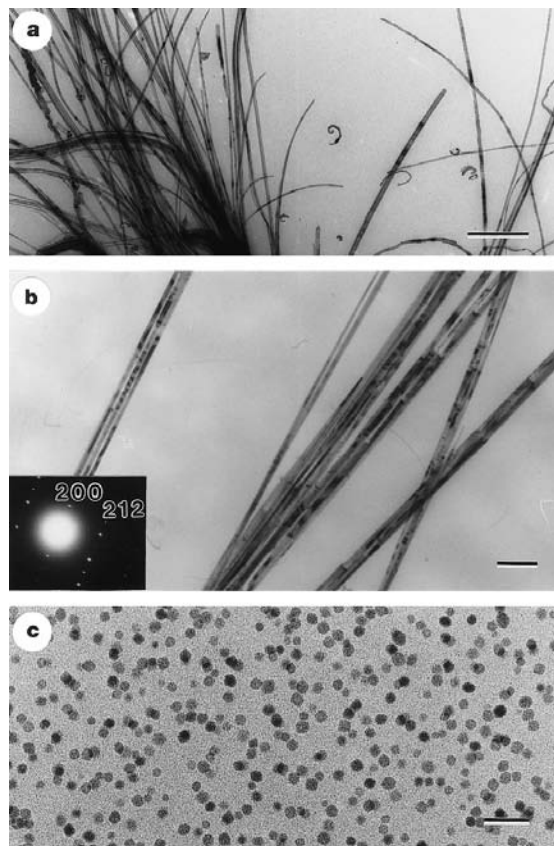


Figure 4 Characterization of BaCrO_4 nanostructures formed at different reactant ratios. **a**, TEM image of bundles of BaCrO_4 nanofilaments prepared in AOT microemulsions at $[\text{Ba}^{2+}] : [\text{CrO}_4^{2-}] \approx 5 : 1$ and $w = 10$; scale bar = 500 nm. **b**, Higher magnification image of single filaments elongated along the crystallographic a axis; scale bar = 200 nm. Inset, $[021]$ zone electron diffraction pattern recorded from an individual fibre. **c**, Spherical BaCrO_4 nanoparticles prepared at $[\text{Ba}^{2+}] : [\text{CrO}_4^{2-}] \approx 1 : 5$ and $w = 10$; scale bar = 50 nm.

($2.7 : 1 \leq [\text{Ba}^{2+}] : [\text{CrO}_4^{2-}] \leq 5.5 : 1$), BaCrO_4 filaments with lengths up to $50 \mu\text{m}$ were deposited in the form of single threads or bundles of loosely aggregated fibres with highly uniform widths, straight edges and well-defined flattened ends (Fig. 4a and b). Similar structures have been reported for BaSO_4 crystallization in microemulsions¹⁹. The filaments were 20–500 nm in width, with an aspect ratio of approximately 1,000. Electron diffraction patterns and lattice images were consistent with single-crystal BaCrO_4 filaments preferentially oriented along the crystallographic a axis. In contrast, when a molar excess of chromate was used ($1 : 4.6 \leq [\text{Ba}^{2+}] : [\text{CrO}_4^{2-}] \leq 1 : 2.7$), no precipitate was observed, and TEM studies showed that the colloidal suspension consisted of cuboidal and spherical BaCrO_4 nanocrystals with a mean size of $11.1 \pm 2.5 \text{ nm}$ (Fig. 4c). When experiments were undertaken at these increased chromate concentrations but with the molar ratio re-established at 1 : 1 by increasing the $\text{Ba}(\text{AOT})_2$ concentration, chains and superlattices, and not discrete spherical nanoparticles, were specifically deposited. This confirmed that the molar ratio, rather than the absolute concentrations, was responsible for the organized structures.

We suggest that on mixing, transfer of water molecules from the chromate-containing microemulsions to the 'dry' $\text{Ba}(\text{AOT})_2$ reverse micelles acts as a strong driving force for the exchange of ions, which results in an increase in intramicellar supersaturation, and nucleation and growth of surfactant-encapsulated BaCrO_4 crystals. At molar equivalence, there is no net charge at the crystal surfaces, so

that the particles develop regular faces and a prismatic morphology in accordance with the unit cell symmetry, and aggregate into ordered chains by the mechanism proposed in Fig. 3. In contrast, a molar excess of CrO_4^{2-} offsets the growth anisotropy, owing to excess negative surface charge, to produce spherical nanoparticles. The AOT molecules are also negatively charged, so there is negligible interaction (adsorption) with the particle surface, and therefore no driving force for surfactant-induced aggregation into longer range structures. An excess of Ba^{2+} ions, on the other hand, results in highly positively charged clusters that interact with the AOT head-groups to such an extent that further growth into regularly shaped nanoparticles is inhibited. Instead, large unstructured micellar aggregates are formed which precipitate from solution and slowly transform over a period of days and weeks into filamentous structures (M.L. and S.M., unpublished data). □

Methods

Sodium AOT was converted to $\text{Ba}(\text{AOT})_2$ by direct precipitation as described elsewhere¹⁹. The following procedure was used for the synthesis of BaCrO_4 within microemulsions. Typically, 0.18 ml of aqueous Na_2CrO_4 solution (0.02–0.50 M, pH ≈ 9) were added with shaking to 10 ml of a solution of NaAOT dissolved in iso-octane (0.1 M) to give yellow suspensions of microemulsion droplets with water to surfactant molar ratio $w = [\text{H}_2\text{O}] : [\text{NaAOT}] = 10$. To introduce Ba^{2+} ions into the reaction, a small amount of $\text{Ba}(\text{AOT})_2$ dissolved in iso-octane (0.395 ml, 0.05 M, $w < 1$) was added to 10 ml of the chromate-containing microemulsion to give a final NaAOT : $\text{Ba}(\text{AOT})_2$ molar ratio of 50 : 1. Under these conditions, the $[\text{Ba}^{2+}] : [\text{CrO}_4^{2-}]$ molar ratio was systematically changed between values of 5.5 : 1 to 1 : 4.6 by modifying the concentration of Na_2CrO_4 in the microemulsion water droplets. Samples for TEM were collected within one week directly from the microemulsion fluid as well as the precipitated material, and air-dried and washed with pure iso-octane on the TEM grids.

Received 20 April; accepted 21 October 1999.

- Petit, C., Taleb, A. & Pileni, M. P. Self-organization of magnetic nanosized cobalt particles. *Adv. Mater.* **10**, 259–261 (1998).
- Taleb, A., Petit, C. & Pileni, M. P. Optical properties of self-assembled 2D and 3D superlattices of silver nanoparticles. *J. Phys. Chem.* **102**, 2214–2220 (1998).
- Taleb, A., Petit, C. & Pileni, M. P. Synthesis of highly monodisperse silver nanoparticles from AOT reverse micelles: A way to 2-D and 3-D self-organization. *Chem. Mater.* **9**, 950–959 (1997).
- Motte, L., Billoudet, F., Lacaze, E. & Pileni, M. P. Self-organization of size-selected nanoparticles into three-dimensional superlattices. *Adv. Mater.* **8**, 1018–1020 (1996).
- Sager, W. F. C. Controlled formation of nanoparticles from microemulsions. *Curr. Opin. Colloid Interf. Sci.* **3**, 276–283 (1998).
- Murray, C. B., Kagan, C. R. & Bawendi, M. G. Self-organization of CdSe nanocrystals into three-dimensional quantum dot superlattices. *Science* **270**, 1335–1338 (1995).
- Vossmeier, T. et al. A double-diamond superlattice built-up of $\text{Cd}_{17}\text{S}_4(\text{SCH}_2\text{OH})_{26}$ clusters. *Science* **267**, 1476–1479 (1995).
- Wang, Z. L. Structural analysis of self-assembling nanocrystal superlattices. *Adv. Mater.* **10**, 13–30 (1998).
- Andres, R. P. et al. Self-assembly of a two-dimensional superlattice of molecularly linked metal clusters. *Science* **273**, 1690–1693 (1996).
- Brust, M., Bethell, D., Schiffrin, D. J. & Kiely, C. J. Novel gold-dithiol nanonetworks with non-metallic electronic properties. *Adv. Mat.* **7**, 795–797 (1995).
- Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382**, 607–609 (1996).
- Alivisatos, A. P. et al. Organization of "nanocrystal molecules" using DNA. *Nature* **382**, 609–611 (1996).
- Li, M., Wong, K. K. W. & Mann, S. Organization of inorganic nanoparticles using biotin-streptavidin connectors. *Chem. Mater.* **11**, 23–26 (1999).
- Shenton, W., Davis, S. A. & Mann, S. Directed self-assembly of nanoparticles into macroscopic materials using antibody-antigen recognition. *Adv. Mater.* **11**, 449–452 (1999).
- Shenton, W., Pum, D., Sleytr, U. B. & Mann, S. Biocrystal templating of CdS superlattices using self-assembled bacterial S-layers. *Nature* **389**, 585–587 (1998).
- Dieluweit, S., Pum, D. & Sleytr, U. B. Formation of a gold superlattice on an S-layer with square lattice symmetry. *Supramol. Sci.* **5**, 15–19 (1998).
- Davis, S. A., Burkett, S. L., Mendelson, N. H. & Mann, S. Bacterial templating of ordered macrostructures in silica and silica-surfactant mesophases. *Nature* **385**, 420–423 (1997).
- Davis, S. A. et al. Brittle bacteria: A biomimetic approach to the formation of fibrous composite materials. *Chem. Mater.* **10**, 2516–2524 (1998).
- Hopwood, J. D. & Mann, S. Synthesis of barium sulfate nanoparticles and nanofilaments in reverse micelles and microemulsions. *Chem. Mater.* **9**, 1819–1828 (1997).

Acknowledgements

We thank S. A. Davis for valuable discussions, the University of Bristol for a postgraduate scholarship to M.L., and the Max-Planck Society for financial support to H.S.

Correspondence and requests for materials should be addressed to S.M. (e-mail: s.mann@bris.ac.uk).

Role of vertical mixing in controlling the oceanic production of dimethyl sulphide

Rafel Simó*† & Carlos Pedrós-Alió*

* Department of Marine Biology and Oceanography, Institut de Ciències del Mar (CSIC), Pg Joan de Borbó s/n, 08039 Barcelona, Catalonia, Spain

† Department of Environmental Chemistry, Institut d'Investigacions Químiques i Ambientals (CSIC), Jordi Girona 18-26, 08034 Barcelona, Catalonia, Spain

Marine microbiota are important for the global biogeochemical sulphur cycle, by making possible the transfer of reduced sulphur from the ocean to the atmosphere in the form of dimethyl sulphide^{1,2}, DMS. Subsequent oxidation of DMS to acidic aerosols influences particle nucleation and growth over the oceans³, and so has the potential to influence radiative balance and global climate. It has been suggested⁴ that this plankton–climate interaction is self-regulated, but tests of this hypothesis have remained elusive as little is known about the feedback effects of climate on the marine DMS cycle⁵. DMS is produced by enzymatic cleavage of the abundant algal component dimethylsulphoniopropionate⁵ (DMSP), which suggests a high potential for DMS generation in the ocean. But there are competing processes⁶ that utilize DMSP in the food web without producing DMS, and the external controls on these processes are unknown. Here we present data of DMSP consumption, DMS production and mixing-layer depths (which are driven by climate) in the subpolar North Atlantic, and compare these data with published results from other latitudes. We find evidence that the mixing-layer depth has a substantial influence on DMS yield in the short term. This finding, combined with the seasonal effect of vertical mixing on plankton succession and food-web structure, suggests that climate-controlled mixing controls DMS production over vast regions of the ocean.

The ACSOE (Atmospheric Chemistry Studies in the Oceanic Environment) North Atlantic Experiment was performed between 6 June and 9 July 1998 in an area around 60° N 21° W (~400 km south of Iceland). Satellite imagery showed that the region was characterized by moderate chlorophyll *a* levels (0.4–2 µg l⁻¹), and patches of high reflectance at a wavelength of 555 nm, indicating the presence of coccolithophorid algae. Biological transformation rates for DMS and DMSP were measured in dark incubations of surface sea water^{7,8}. DMSP consumption and one of its consequences, DMS production, were not always coupled⁷. On the contrary, couplings and de-couplings occurred both in time and space. Thus, DMS production represented between 5% and 100% of DMSP consumption, changing from site to site in the same area, and from day to day in the same water mass⁷.

We found a striking relationship between the DMS yield of DMSP consumption and the mixing-layer depth (MLD, Fig. 1). This relationship consisted of an asymmetric curve with a minimum yield at MLDs of 15–20 m, higher yields at deeper mixing depths, and yields close to 100% in very shallow mixed layers. The few data on DMS yields available in the literature fitted extremely well into the general pattern set by the Atlantic data (Fig. 1). If the whole data set is analysed by splitting the curve into two straight lines with different and opposite-sign slopes, each regression explains about 80% of the variance of its corresponding data range. We note that the data assembled were obtained by different methods^{8–11}, in non-coastal areas (>20 m depth), at different latitudes of the Northern Hemisphere (between 5° and 60° N), in different seasons (between February and October), and that they covered a wide range of seawater temperatures (8–26 °C). Therefore, their fit into the same

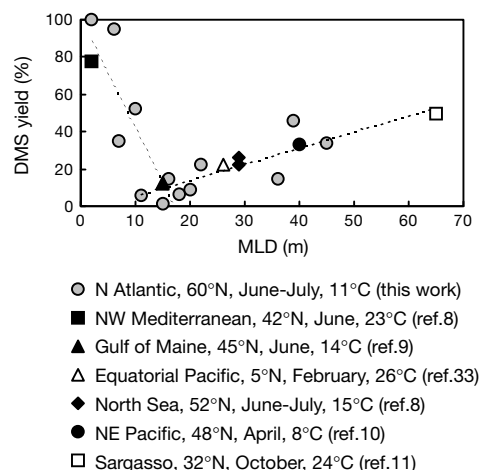


Figure 1 DMS production and vertical mixing. Shown is the DMS production yield from DMSP, that is (DMS production rate/DMSP consumption rate) × 100, versus mixing layer depth (MLD, defined as the depth at which σ_t increased by 0.03 kg m⁻³ or temperature dropped by 0.1 °C). Data are shown for the subpolar North Atlantic at 60° N, 21° W (grey circles; this work) and other oceanic waters (other symbols). The other data have been extracted from the literature (refs 8, 9, 10, 11, 33) with help from the authors in some cases. In the North Atlantic samples, calculation of error propagation resulted in coefficients of variation for yield estimates of 18–39%. The regression equations obtained (dotted lines) are:

$$\text{DMS yield} = (-6\text{MLD}) + 100 \quad r^2 = 0.78 \quad (\text{MLD} < 20 \text{ m})$$

$$\text{DMS yield} = (0.9\text{MLD}) - 3 \quad r^2 = 0.78 \quad (\text{MLD} > 10 \text{ m})$$

There are a further three data points not included in this graph: they all belong to the Sargasso Sea, 32° N, January, 19 °C. The MLD was ~250 m, and the DMS yields were 11%, 14% and 21% (data from ref. 11), thus suggesting that the trend of the right half of the curve cannot be extrapolated to deep mixing conditions, which favour non-DMS-producing, DMSP consumption pathways.

curve strongly suggests that this relationship reflects a robust pattern common to a large portion of the ocean.

Is there an explanation for such an unusual pattern? DMSP is an intracellular biosynthate in phytoplankton. It is released from algal cells by exudation and through grazing, autolysis, bacterial and viral attack, and thus becomes available to enzymatic degradation or assimilation⁵. The enzymes that cleave DMSP to DMS (DMSP-lyases) have been identified in both algal and bacterial species¹². Algal DMSP-lyase is segregated from intracellular DMSP, and acts either extracellularly upon DMSP exudation¹³ or when substrate and enzyme mix together upon cell disruption by lysis or grazing¹⁴. Bacteria use DMSP-lyase to cleave the DMSP released by phytoplankton⁶ either free in solution or closely attached to algal cells and aggregates. Pathways for DMSP utilization other than cleavage have been observed only in bacteria and, to a lesser extent, in micrograzers¹². In the past few years, evidence has been accumulating showing that a large fraction of DMSP is demethylated and demethylated by bacterioplankton to produce methanethiol^{6,9}, with methanethiol being used for methionine and eventually protein synthesis¹⁵. Recent work suggests that DMSP assimilation alone may satisfy about 100% of the sulphur demand of marine heterotrophic bacteria (R. P. Kiene *et al.* and R.S. *et al.*, unpublished results).

There are therefore two families of pathways for DMSP degradation: one leads to DMS and is carried out by both algal and bacterial enzymes (DMSP cleavage), and another one does not lead to DMS and is carried out essentially only by bacteria (DMSP assimilation). Within this picture, any factor that affects differentially algal and bacterial activities and turnovers may lead to changes in the relative

importance of the two families of pathways and, therefore, may result in different DMS yields from DMSP.

Recent studies have shown that vertical mixing affects both primary and heterotrophic bacterial production by subjecting microorganisms to different degrees of exposure to ultraviolet (UV-B) and total solar radiation^{16–18}. At the same latitude of the Atlantic waters studied here (but in the Southern Ocean), photo-inhibition of photosynthesis was strongest when the water column was mixed to about 20 m depth. In contrast photoinhibition was less at deeper MLD, and photosynthesis was closest to its maximal potential values in very shallow mixing layers (MLD < 10 m)¹⁶, thus configuring an asymmetric curve of effective/potential photosynthesis against MLD which is remarkably similar to our curve of Fig. 1. Other authors have shown that heterotrophic bacterial activity is most inhibited by UV-B when kept in very shallow waters (down to 10 m), whereas deeper mixing allows DNA repair and recovers activity^{17,18}. Moreover, bacteria are more sensitive to photo-damage than phytoplankton¹⁷. Hence, shallow mixing layers would be less favourable to bacterioplankton: higher exposure to UV-B would depress heterotrophic production^{17,18} and would lower bacterial sulphur demand. This is consistent with the higher DMSP cleavage (DMS production) found at shallow mixing depths. Relaxation of UV-inhibition by deeper mixing would increase bacterial sulphur demand, thus enlarging the contribution of non-DMS-producing assimilative pathways to DMSP consumption. The occurrence of a minimum of the DMS yield at intermediate mixing depths (down to 15–20 m), and the overall similarity of the curve in Fig. 1 to that of photosynthetic activity (ref. 16), suggests that the direct role of phytoplankton in DMSP-to-DMS conversion is larger than currently recognized. In this context, recent reports have shown that the coccolithophore *Emiliania huxleyi* may produce microgradients of acrylate and DMS around the cell without the need to be ingested and lysed¹⁴. Also, in waters dominated by the flagellate haptophyte *Phaeocystis* sp., DMS production by algal DMSP-lyase accounts for a very significant fraction of the annual DMS production¹⁹. Therefore, it is plausible that differential photophysical effects of the exudation/uptake rates and enzymatic activities of phytoplankton and heterotrophic bacteria can explain the pattern in Fig. 1. Regardless of the actual explanation, the results shown in Fig. 1 indicate that mixing exerts a profound effect on DMS production in the short term.

Vertical mixing probably also affects DMS production at longer timescales. At temperate and subpolar latitudes, seasonal changes in the stratification regime are thought to determine not only the seasonality of primary production, but also the phytoplankton succession. Thus, winter and early spring mixed waters are characterized by blooms of large diatoms and high chlorophyll *a* concentrations (Fig. 2a and b), that is, high biomass but low potential for DMS production, as most diatoms are weak DMSP producers⁵. The stratification of surface waters during spring lowers phytoplankton biomass (Fig. 2a and b) but favours the development of small flagellates, coccolithophores and dinoflagellates²⁰, most of them strong DMSP producers⁵. Moreover, phytoplankton succession is accompanied by profound seasonal changes in food-web structure: in mixed waters during winter–spring, autotrophs and grazers are larger, and part of the primary production sinks and swims out of the photic layer, whereas in strongly stratified waters most of the primary production in the mixed layer is recycled *in situ* through the microbial loop²¹, thus facilitating a more rapid turnover of phytoplanktonic DMSP²².

This pattern of phytoplankton succession and changes in the food-web structure should favour an increase in DMS production as water stratifies, peaking along with the biomass of DMSP producers in spring. However, in a recently published, open-ocean temporal series for the Sargasso Sea²², highest DMS concentrations in surface waters were not coincident with the highest DMSP concentrations around May (Fig. 2d). Instead, the highest DMS concentrations

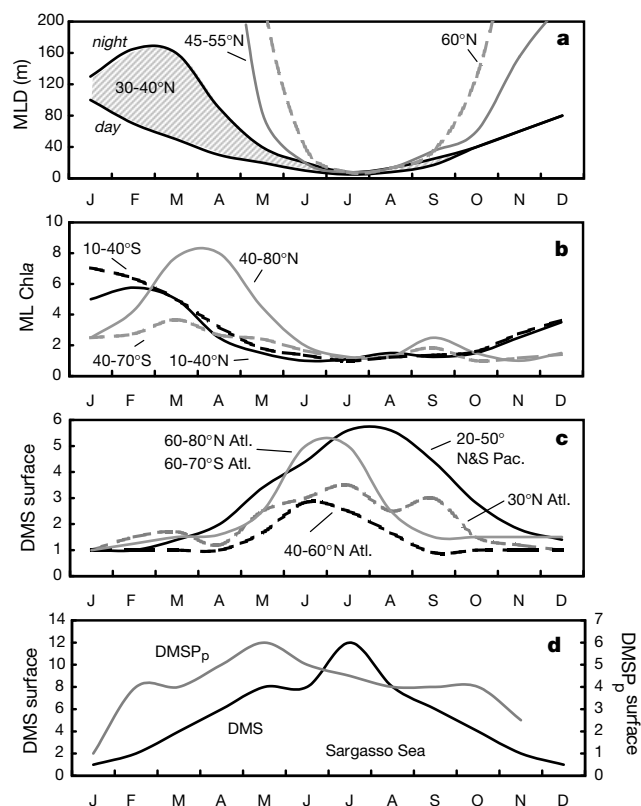


Figure 2 The ‘summer DMS paradox’ and vertical mixing in the surface ocean. Shown is the seasonality of vertical mixing, chlorophyll *a* and DMS concentrations in surface waters at a broad range of latitudes in the world oceans. **a**, Monthly variation in MLD. The lines labelled 60°N and 45–55°N represent daily averages. The shaded area represents the amplitude of the day/night variability of vertical mixing at the latitude range indicated (after refs 34, 35). **b**, Monthly variation in the chlorophyll *a* concentration in the mixed layer in four different latitudinal regions (after ref. 34). Southern Hemisphere data have been shifted by 6 months. Values are normalized with respect to the annual minimum. **c**, Monthly variation in surface DMS concentrations in diverse latitudinal regions (after refs 23, 26). Southern Hemisphere data have been shifted by 6 months. Values are normalized with respect to the annual minimum. **d**, Monthly variation in surface particulate DMSP (DMSP_p) and DMS concentration in the Sargasso Sea (from ref. 22). Values are normalized with respect to the annual minimum.

appeared 2 months later, in mid-summer. This lag cannot be explained simply by DMS accumulation over time, as turnover times of DMS in sea water are generally of the order of 0.5 to 3 days (refs 7, 8, 11).

Highest annual concentrations of DMS in surface sea water of a broad range of subtropical, temperate and subpolar open-ocean areas are also found in summer (see, for example, refs 22–25, and compilation in ref. 26), when surface chlorophyll *a* levels are lowest. This can be regarded as the ‘summer DMS paradox’, and is shown in Fig. 2b and c. At latitudes between 10° and 70°, mixed-layer chlorophyll *a* concentrations reach their maximum in late winter and early spring. Conversely, surface DMS generally increases from early spring with the onset of stratification, peaking in summer with concentrations higher than those of winter by a factor of 3–10. As a result, DMS/chlorophyll *a* ratios exhibit a similar pattern, but with a higher summer/winter factor of 10–50.

The observations presented in Figs 1 and 2, along with recent evidence that the DMS yield exhibits short-term variability of the order of days (similar to that of meteorological forcing factors)⁷, point to the existence of some factor able to control the efficiency of DMSP conversion to DMS at both the short⁷ and the seasonal^{11,22} timescales. We suggest that this factor is vertical mixing. In the short term, photophysical effects on the activities of planktonic autotrophs

and heterotrophs, influenced by vertical mixing, would control both the DMS yield and DMS consumption, thus causing DMS accumulation in shallow mixed waters. Short-term and seasonal effects of vertical mixing cooperate to produce maximum DMS concentrations in highly stratified surface waters in summer (Fig. 2), by (1) plankton succession toward stronger DMSP producers and higher contribution of the microbial loop to DMSP turnover, (2) higher efficiency of DMSP-to-DMS conversion (Fig. 1) because of photophysical effects, and probably also nutrient limitation of bacterial sulphur demand, and (3) possible inhibition of bacterial DMS consumption by greater UV-B radiation. What about the photochemical sink for DMS? Greater solar radiation intensities, longer days and shallower mixing depths would make enhanced DMS photochemical losses in summer expected²². However, DMS photo-oxidation is a photosensitized reaction²⁷, and chromophoric dissolved organic matter in the mixed layer is at a minimum in summer²⁸. As a result, effective photochemical loss rates do not seem to be enough to counteract the combination of factors for enhanced DMS production, thus leading to the pattern in Fig. 2.

We can now speculate on the implications of these findings for the hypothesis of a link between plankton and climate proposed by Charlson *et al.*⁴ (also known as the CLAW hypothesis after the initials of the authors). Evidence suggests that most of the sulphur in radiation-reflecting aerosols over the remote ocean is biogenic and comes from the atmospheric oxidation of DMS^{2,3}. The potential of planktonic DMS to affect the radiative balance and climate seems well established. However, tests of whether this constitutes a self-regulated plankton–climate interaction, as suggested by Charlson *et al.*⁴, have remained elusive, essentially because little is known about the feedback effects of climate on marine DMS production. Our findings provide new arguments in support of the feedback hypothesis. First, in the short term, a local increase in solar radiation incident on already stratified waters would further shoal the surface mixing layer, and would cause stimulation of DMS production from the DMSP stock in the plankton assemblage, and DMS accumulation, as required for the feedback hypothesis. Second, on a seasonal scale, highest DMS concentrations would occur in summer in response to maximal solar radiation (Fig. 2c). Increased DMS fluxes to the atmosphere in summer are to be expected in accordance with higher concentrations. This is in agreement with the linear relation found between DMS flux and incident solar radiation for a broad range of latitudinal regions²⁹, which was cited as a supporting argument for the hypothesis of Charlson *et al.*

In the context of longer-term global environmental change, continuous warming of surface sea water might also cause an increase in DMS yield. Current models of CO₂-derived global warming predict stronger and expanded surface stratification of the oceans³⁰; that is, shallower mixed and mixing layers over larger parts of the oceans and over longer seasons of the year. These conditions might favour higher production and higher concentrations of DMS in the open ocean, through the mechanisms described above. Higher aqueous DMS levels in warmer waters would involve larger emissions to the atmosphere, which would have climate-cooling effects inverse to those of the CO₂ build up. Therefore, in the context of global warming, a decrease in atmospheric CO₂ uptake by the oceans^{30,31} might go along with a larger release of DMS from open-ocean waters at middle and subpolar latitude. Quantifying this coupled effect represents a formidable challenge, especially as the response to mixing of DMS production in the marine regions not covered by our data—such as equatorial, polar, and coastal productive waters—remains unknown.

We have found a link between a climate-driven parameter (MLD) and DMS production in the open ocean. Moreover, this link lies in the direction of the negative feedback suggested by Charlson *et al.*⁴ 12 years ago. Most previous attempts to address this feedback have focused on the physiological effects of light and temperature on

DMSP biosynthesis in phytoplankton, but gave unconvincing results⁵. Our systemic approach indicates that it is the set-up of the overall DMS-producing and DMS-consuming plankton community, and the relative activities within it, that is sensitive to local weather and to climate. This approach also shows that a sulphur-mediated negative feedback between marine plankton and climate (both in the short and the long term) can operate through the profound effects of MLD on the biogeochemistry of the surface ocean^{16,17,32}. □

Received 10 June; accepted 18 October 1999.

- Lovelock, J. E., Maggs, R. J. & Rasmussen, R. A. Atmospheric sulphur and the natural sulphur cycle. *Nature* **237**, 452–453 (1972).
- Andreae, M. O. & Crutzen, P. J. Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry. *Science* **276**, 1052–1058 (1997).
- Clarke, A. D. *et al.* Particle nucleation in the tropical boundary layer and its coupling to marine sulfur sources. *Science* **282**, 89–92 (1998).
- Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**, 655–661 (1987).
- Malin, G. & Kirst, G. O. Algal production of dimethyl sulfide and its atmospheric role. *J. Phycol.* **33**, 889–896 (1997).
- Taylor, B. F. & Visscher, P. T. in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (eds Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O.) 265–276 (Plenum, New York, 1996).
- Simó, R. & Pedrós-Alió, C. Short-term variability in the open ocean cycle of dimethylsulfide. *Glob. Biogeochem. Cycles* **13**, 1173–1181 (1999).
- Simó, R., Pedrós-Alió, C., Malin, G. & Grimalt, J. O. Biological cycling of DMS, DMSP and DMSP in contrasting marine waters. *Limnol. Oceanogr.* (submitted).
- Kiene, R. P. Production of methane thiol from dimethylsulfoniopropionate in marine surface waters. *Mar. Chem.* **54**, 69–83 (1996).
- Bates, T. S. *et al.* The cycling of sulfur in surface seawater of the northeast Pacific. *J. Geophys. Res.* **99**, 7835–7843 (1994).
- Ledyard, K. M. & Dacey, J. W. H. Microbial cycling of DMSP and DMS in coastal and oligotrophic seawater. *Limnol. Oceanogr.* **41**, 33–40 (1996).
- Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O. (eds) *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (Plenum, New York, 1996).
- Stefels, J. & van Boekel, W. H. M. Production of DMS from dissolved DMSP in axenic cultures of the marine phytoplankton species *Phaeocystis* sp. *Mar. Ecol. Prog. Ser.* **97**, 11–18 (1993).
- Wolfe, G. V., Steinke, M. & Kirst, G. O. Grazing-activated chemical defence in a unicellular marine alga. *Nature* **387**, 894–897 (1997).
- Kiene, R. P., Linn, L. J., González, J., Moran, M. A. & Bruton, J. A. Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl. Environ. Microbiol.* **65**, 4549–4558 (1999).
- Neale, P. J., Davis, R. F. & Cullen, J. J. Interactive effects of ozone depletion and vertical mixing on photosynthesis of Antarctic phytoplankton. *Nature* **392**, 585–589 (1998).
- Herndl, G. J., Müller-Niklas, G. & Frick, J. Major role of ultraviolet-B in controlling bacterioplankton growth in the surface layer of the ocean. *Nature* **361**, 717–719 (1993).
- Jeffrey, W. H. *et al.* Diel and depth profiles of DNA photodamage in bacterioplankton exposed to ambient solar ultraviolet radiation. *Mar. Ecol. Prog. Ser.* **137**, 283–291 (1996).
- van den Berg, A. J., Turner, S. M., van Duyl, F. C. & Ruardij, P. Model structure and analysis of dimethylsulphide (DMS) production in the North Sea, considering phytoplankton dimethylsulphoniopropionate- (DMSP) lyase and eutrophication effects. *Mar. Ecol. Prog. Ser.* **145**, 233–244 (1996).
- Margalef, R. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol. Acta* **1**, 493–509 (1978).
- Kiorboe, T. Turbulence, phytoplankton cell size, and the structure of pelagic food webs. *Adv. Mar. Biol.* **29**, 1–72 (1993).
- Dacey, J. W. H., Howse, F. A., Michaels, A. F. & Wakeham, S. G. Temporal variability of dimethylsulfide and dimethylsulfoniopropionate in the Sargasso Sea. *Deep-Sea Res.* **145**, 2085–2104 (1998).
- Bates, T. S., Cline, J. D., Gammon, R. H. & Kelly-Hanse, S. R. Regional and seasonal variations in the flux of oceanic dimethylsulphide to the atmosphere. *J. Geophys. Res.* **92**, 2930–2938 (1987).
- Nguyen, B. C., Mihalopoulos, N. & Belviso, S. Seasonal variation of atmospheric dimethylsulphide at Amsterdam Island in the southern Indian Ocean. *J. Atmos. Chem.* **11**, 123–141 (1990).
- Tarrasón, L., Turner, S. & Floisand, I. Estimation of seasonal dimethyl sulphide fluxes over the North Atlantic Ocean and their contribution to European pollution levels. *J. Geophys. Res.* **100**, 11623–11639 (1995).
- Kettle, A. J. *et al.* A global database of sea surface dimethylsulphide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Glob. Biogeochem. Cycles* **13**, 399–444 (1999).
- Brimblecombe, P. & Shooter, D. Photo-oxidation of dimethylsulphide in aqueous solution. *Mar. Chem.* **19**, 343–353 (1986).
- Siegel, D. A. & Michaels, D. F. Quantification of non-algal attenuation in the Sargasso Sea: implications for biogeochemistry and remote sensing. *Deep-Sea Res.* **43**, 321–345 (1996).
- Bates, T. S., Charlson, R. J. & Gammon, R. H. Evidence for the climatic role of marine biogenic sulphur. *Nature* **329**, 319–321 (1987).
- Sarmiento, J. L., Hughes, T. M. C., Stouffer, R. J. & Manabe, S. Simulated response of the ocean carbon cycle to anthropogenic climate warming. *Nature* **393**, 245–249 (1998).
- Joos, F., Plattner, G.-K., Stocker, T. F., Marchal, O. & Schmitter, A. Global warming and marine carbon cycle feedbacks on future atmospheric CO₂. *Science* **284**, 464–467 (1999).
- Karl, D. M. *et al.* Ecosystem changes in the North Pacific subtropical gyre attributed to the 1991–92 El Niño. *Nature* **373**, 230–234 (1995).
- Kiene, R. P. & Bates, T. S. Biological removal of dimethyl sulphide from sea water. *Nature* **345**, 702–705 (1990).

34. Falkowski, P. G. & Raven, J. A. *Aquatic Photosynthesis* (Blackwell Science, Malden, 1997).
 35. Woods, J. D., Barkmann, W. & Strass, V. in *The Ocean Surface* (eds Toba, Y. & Mitsuyasu, H.) 487–507 (Reidel, Dordrecht, 1985).

Acknowledgements

We thank all scientists, officers and crew aboard the RRS *Discovery* for assistance, P. Liss and T. Jinkells for leadership during fieldwork, W. Broadgate for setting up the cruise, P. Machine and G. Monocoffé of the BODC for CTD and accessory data management, J. Grimalt for analytical facilities, and R. Kiene and K. Ledyard for providing accessory information to published data. R. Kiene also provided valuable comments. This work was supported by the Spanish CICYT through the Programa Nacional de Ciencia y Tecnología Marinas (CyTMAR) and the British NERC through the Atmospheric Chemistry Studies in the Oceanic Environment (ACSOE) Community Programme.

Correspondence and requests for materials should be addressed to R.S.
 (e-mail: rsimo@icm.csic.es).

The effect of climate change on ozone depletion through changes in stratospheric water vapour

Daniel B. Kirk-Davidoff, Eric J. Hints, James G. Anderson & David W. Keith

Department of Earth and Planetary Sciences and Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA

Several studies have predicted substantial increases in Arctic ozone depletion due to the stratospheric cooling induced by increasing atmospheric CO₂ concentrations^{1,2}. But climate change may additionally influence Arctic ozone depletion through changes in the water vapour cycle. Here we investigate this possibility by combining predictions of tropical tropopause temperatures from a general circulation model with results from a one-dimensional radiative convective model, recent progress in understanding the stratospheric water vapour budget, modelling of heterogeneous reaction rates and the results of a general circulation model on the radiative effect of increased water vapour³. Whereas most of the stratosphere will cool as greenhouse-gas concentrations increase, the tropical tropopause may become warmer, resulting in an increase of the mean saturation mixing ratio of water vapour and hence an increased transport of water vapour from the troposphere to the stratosphere. Stratospheric water vapour concentration in the polar regions determines both the critical temperature below which heterogeneous reactions on cold aerosols become important (the mechanism driving enhanced ozone depletion) and the temperature of the Arctic vortex itself. Our results indicate that ozone loss in the later winter and spring Arctic vortex depends critically on water vapour variations which are forced by sea surface temperature changes in the tropics. This potentially important effect has not been taken into account in previous scenarios of Arctic ozone loss under climate change conditions.

The winter/spring depletion of ozone in the polar vortex is determined by (1) halogen activation, (2) sunlight (which drives the photochemical reactions that catalytically destroy ozone) and (3) deactivation of chlorine to reservoir species. Because of the dependence on insolation, ozone is most sensitive to active chlorine in the spring. Halogens are activated—converted to free-radical form—in nearly every Arctic winter, when temperatures fall below about 195 K. Deactivation occurs at different times each year, depending on stratospheric temperatures and dynamics. For example, the 1996–97 winter had particularly severe ozone depletion because the Arctic vortex stayed cold later, even though

temperatures were not particularly low⁴. Increases in the concentration of water vapour in the atmosphere raise the threshold temperature for halogen activation, thus allowing polar processing to continue later in the season and maintaining high concentrations of ClO. In a warm winter with temperatures near the threshold for chlorine activation, increases in water vapour may allow activation to occur when it would otherwise not. Thus, ozone losses will not simply scale with chlorine loading in the stratosphere in coming decades.

Previous studies of the relationship between climate change and ozone depletion have focused on the influence of increasing carbon dioxide on polar stratospheric temperatures, and thus on heterogeneous ozone chemistry^{1,2}. Recent progress in the observation and understanding of the stratospheric water vapour budget^{5,6} now makes possible an evaluation of the chain linking greenhouse-gas increases to warming in the tropics, increased water vapour in the stratosphere, increased aerosol activity in the polar vortices, and thus increased high-latitude ozone depletion. Because ozone depletion depends so strongly on the presence of temperatures low enough to initiate heterogeneous chemistry, we adopt $(T_\alpha - T_v)$ as a predictor of ozone loss, where T_v is the temperature of the polar vortex, and T_α is the temperature below which heterogeneous chemistry on aerosols becomes important, yielding dramatic increases in the concentration of ClO. We can express the rest of this chain mathematically as follows:

$$\frac{\partial(T_\alpha - T_v)}{\partial T_s} = \frac{\partial(T_\alpha - T_v)}{\partial [H_2O]_v} \frac{\partial [H_2O]_v}{\partial [H_2O]_e} \frac{\partial [H_2O]_e}{\partial T_e} \frac{\partial T_e}{\partial T_s} \quad (1)$$

T_e is the temperature of the tropical tropopause, T_s is the surface temperature in the tropics, $[H_2O]_e$ is the concentration of water vapour entering the tropical stratosphere, and $[H_2O]_v$ is the concentration of water vapour in the polar vortices. We treat here only those variations in T_e due to changes in the tropical climate (and reflected in changes in T_s), thus neglecting changes in T_e due, for instance, to changes in stratospheric circulation. We have also neglected the terms $\partial T_e / \partial p_e$ and $\partial [H_2O]_e / \partial p_e$, where p_e is the pressure at which air enters the stratosphere (that is, the pressure of the tropical tropopause), because, as we will discuss, we believe $\partial p_e / \partial T_s$, the dependence of tropical tropopause pressure on tropical surface temperature, to be small.

We begin by considering $\partial(T_\alpha - T_v) / \partial [H_2O]_v$. To evaluate this term we must analyse the dependence of ozone depletion on temperature and water vapour. To a good approximation, the time rate of change of ozone in the polar vortices is given by the sum of the rate-limiting steps of catalytic cycles involving ClO and

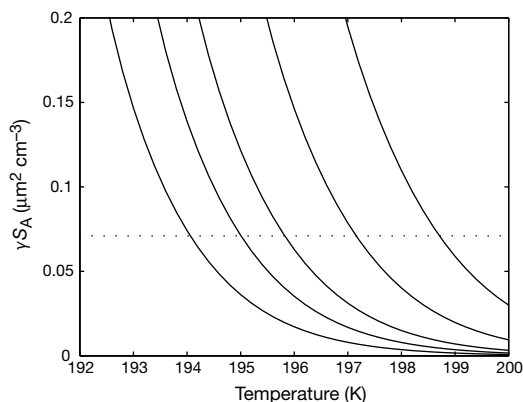


Figure 1 Temperature and water dependence of γS_A at 50 mbar. S_A is the aerosol surface area, and γ is the reaction probability per collision for the reaction $HCl + ClONO_2$. Curves are plotted for water vapour concentrations of 5, 6, 7, 9 and 12 p.p.m.v., from left to right. $\partial T_\alpha / \partial [H_2O]_v$ is evaluated along the dotted line, which indicates the value of γS_A at the temperature (195 K) at which heterogeneous ozone depletion is observed to become significant for the present water vapour mixing ratio of ~6 p.p.m.v.

BrO (ref. 7).

$$\frac{d[\text{O}_3]}{dt} = -2k_{\text{ClO}-\text{ClO}}[\text{ClO}]^2 - 2k_{\text{ClO}-\text{BrO}}[\text{ClO}][\text{BrO}] \quad (2)$$

The total ozone loss is the integral of this loss rate over the volume of the region where ClO is elevated. In the Arctic vortex, unlike the Antarctic, destruction of ozone is never complete, so that both the loss rate and the volume and duration of the region in which ozone loss occurs are important⁸.

The concentration of ClO depends on the partitioning of total inorganic chlorine between ClO and the reservoir species HCl and ClONO₂. In the polar vortices, this partitioning is controlled by heterogeneous reactions of the reservoir species on surfaces, yielding ClO. The quadratic dependence on [ClO] means that the ozone loss rate is very sensitive to increases of both $\gamma_{\text{HCl+ClONO}_2}$, the reaction rate of gaseous ClONO₂ with HCl on aerosols, and S_A , the aerosol surface area per unit volume, as temperature decreases or water vapour increases⁹. Denitrification is also important, and is expected to enhance ozone depletion in a future atmosphere with lower chlorine loading and lower temperatures (and increased water vapour) by limiting deactivation of ClO to ClONO₂ (ref. 10).

Figure 1 shows curves of γS_A plotted versus T_v for various values of $[\text{H}_2\text{O}]_v$. As $[\text{H}_2\text{O}]_v$ increases, so does the temperature (T_α) at which γS_A rises abruptly. The temperature–water vapour dependence of S_A is taken from ref. 11, and that of γ from ref. 12. These relationships have been substantially verified by aircraft measurements of ClO (ref. 13).

Using the data presented in Fig. 1, we can evaluate $\partial T_\alpha/\partial[\text{H}_2\text{O}]_v$ for a range of water vapour mixing ratios. We note that in the present conditions, $[\text{H}_2\text{O}]_v$ is ~ 6 p.p.m.v. (ref. 14), and that heterogeneous ozone depletion chemistry becomes important at temperatures below 195 K (at 50 mbar). We thus choose $\gamma S_A = 0.071$ as the value of γS_A corresponding to T_α . We can then set $\partial T_\alpha/\partial[\text{H}_2\text{O}]_v = \partial T/\partial[\text{H}_2\text{O}]$ evaluated along the parametric curve $\gamma S_A(T, [\text{H}_2\text{O}]) = c$, where $c = 0.071$. For present conditions, $\partial T_\alpha/\partial[\text{H}_2\text{O}]_v = 0.8$ K p.p.m.v.⁻¹. This value is insensitive to variation in c , but decreases logarithmically with increasing $[\text{H}_2\text{O}]_v$, falling to 0.7 K p.p.m.v.⁻¹ for $[\text{H}_2\text{O}]_v = 8$ p.p.m.v.

Forster and Shine have recently evaluated $\partial T_v/\partial[\text{H}_2\text{O}]_v$ (ref. 3). Using both a general circulation model (GCM) and a column model, they found that a uniform increase of 0.7 p.p.m.v. of water vapour in the stratosphere resulted in a cooling of the polar vortices by 4–6 K in the spring, implying that $\partial T_v/\partial[\text{H}_2\text{O}]_v$ ranges from 6 to 9 K p.p.m.v.⁻¹. Thus, the radiative effect of an increase in water vapour on vortex temperature dwarfs the chemical effect of such an increase discussed above. They point out, however, that the net cooling effect of observed changes in stratospheric CO₂, H₂O and O₃ over-predicts the observed stratospheric temperature trends by a factor of 2 or more. Thus, either the stratospheric water vapour trend or stratospheric climate modelling has large errors.

We next consider $\partial[\text{H}_2\text{O}]_v/\partial[\text{H}_2\text{O}]_e$. Air reaches the polar vortices along trajectories high in the stratosphere, without mixing with air below the 380 K potential temperature surface¹⁵. Thus water vapour concentrations in the polar vortices are governed predominantly by the concentration of water vapour and CH₄ entering the stratosphere in the tropics, and $\partial[\text{H}_2\text{O}]_v/\partial[\text{H}_2\text{O}]_e$ is therefore ~ 1.0 . Oxidation of CH₄ adds about 2 p.p.m.v. of water vapour to air in the polar vortices¹⁴ but has an insignificant effect on $\partial[\text{H}_2\text{O}]_v/\partial[\text{H}_2\text{O}]_e$.

Recent improvements in our understanding of the tropical tropopause indicate that $\partial[\text{H}_2\text{O}]_e/\partial T_e$ can be evaluated to useful accuracy simply by assuming that $[\text{H}_2\text{O}]_e = [\text{H}_2\text{O}]^*(T_e, p_e)$, where $[\text{H}_2\text{O}]^*(T, p)$ is the saturation mixing ratio of water vapour as a function of temperature and pressure. T_e and p_e are the temperature and pressure at the tropopause, defined as the altitude of minimum temperature, and the bar represents an average over all tropopause points in the range 10° S to 10° N. A careful examination of radiosonde temperature measurements along with a survey of relevant

stratospheric water vapour and CH₄ measurements⁵ has shown that $[\text{H}_2\text{O}]^*(T_e, p_e)$, calculated from radiosonde data, is consistent with observations of stratospheric water vapour. Furthermore, a Green's function analysis of water vapour and CO₂ profiles shows that seasonal variations of tropical stratospheric water vapour are consistent with variations in the mean of daily tropopause saturation mixing ratios in sounding data¹⁶. This assumption allows us to calculate $\partial[\text{H}_2\text{O}]_e/\partial T_e$ using the Clausius–Clapeyron equation. Present tropopause mean temperature and pressure of 191 K and 95 mbar imply a value of 0.8 p.p.m.v. K⁻¹. $\partial[\text{H}_2\text{O}]_e/\partial T_e$ itself increases with temperature, rising to 1.0 p.p.m.v. K⁻¹ at 193 K, and 1.4 p.p.m.v. K⁻¹ at 195 K, assuming p_e remains constant.

To estimate the value of $\partial T_e/\partial T_s$, we turn to modelling studies of the structure of the atmosphere's response to increased greenhouse-gas loading. We first consider output of the National Center for Atmospheric Research (NCAR) CCM3 atmospheric model¹⁷ coupled with a slab ocean model, and run to equilibrium at pre-industrial and at doubled CO₂ concentrations (J. T. Kiehl, personal communication). The temperature difference between the two equilibria (Fig. 2) increases from the surface upwards through the upper troposphere, since the convection parametrization tends to adjust the tropical atmosphere towards a moist-adiabatic lapse rate. In the region of the tropopause, the temperature difference abruptly shifts from large positive values to large negative values, yielding a change of 1.0 K in T_e for a change of 1.5 K in T_s . However, the model's low resolution means that it cannot predict changes in tropopause height which might accompany a warming of the tropics.

To evaluate the possible changes in the height of the tropopause, we investigated the behaviour of a one-dimensional radiative–convective model^{18,19} with enhanced resolution (2.5 mbar) in the neighbourhood of the tropopause, and a cooling term of 0.5 K d⁻¹ added in the stratosphere and upper troposphere to represent the Brewer–Dobson circulation. In this model, the height of the tropopause remains constant under a wide variety of climate-change scenarios, lending support to the use of the GCM to predict tropopause temperature changes. However, the temperature change at the tropopause in this model is generally larger than in GCMs for the same change in surface temperature: 2.4 K for a 1.5 K warming at the surface. Here we use the NCAR CCM3 prediction as the conservative basis for our estimate of the change in tropical tropopause temperature and saturation water vapour pressure, and thus of the change in polar vortex water vapour in a

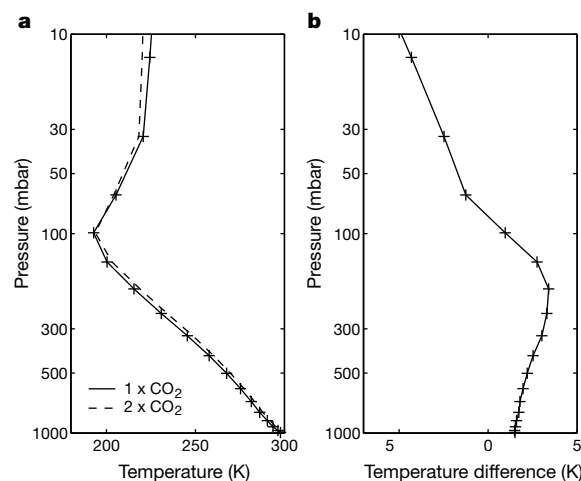


Figure 2 GCM simulation of tropical vertical temperature profile. **a**, Tropical mean (10°S–10°N) temperatures from NCAR CCM3 atmospheric model coupled with a slab ocean, and run to equilibrium with pre-industrial (1 × CO₂, solid line) and doubled (2 × CO₂, dashed line) carbon dioxide concentrations. **b**, Tropical mean temperature difference (2 × CO₂ – 1 × CO₂).

Table 1 Terms relating ozone depletion to greenhouse forcing

Derivative	Estimated value	References
$\frac{\partial T_v}{\partial [\text{H}_2\text{O}]_v}$	−6 to −9 K p.p.m.v. ^{−1}	3
$\frac{\partial T_\alpha}{\partial [\text{H}_2\text{O}]_v}$	0.8 K p.p.m.v. ^{−1}	
$\frac{\partial [\text{H}_2\text{O}]_v}{\partial [\text{H}_2\text{O}]_a}$	1.0	25, 13
$\frac{\partial [\text{H}_2\text{O}]_a}{\partial T_o}$	0.8 p.p.m.v. K ^{−1}	5
$\frac{\partial T_o}{\partial T_s}$	0.7 (GCMs)–1.6 (1-D model)	
dT_s	1–4 K (W ^{−1} m ²)	20

doubled-CO₂ climate. Another GCM, the Goddard Institute for Space Studies (GISS) atmospheric model²⁰, predicts a similar ratio of tropopause warming to surface warming, with equilibrium tropical tropopause warming of 2.0 K for 3.1 K at the surface in a run with doubled CO₂ (R. Ruedy, personal communication). Thus, we estimate $\partial T_o/\partial T_s$ to be 0.7. Use of the high-resolution one-dimensional radiative convective model would yield an estimate of 1.6.

We now estimate the effect of climate change on ozone depletion via water vapour changes. Table 1 lists estimated values and references for each of the derivatives in equation (1). Multiplying these out, and assuming a relatively small value for $\partial T_v/\partial [\text{H}_2\text{O}]_v$ of −6 K p.p.m.v.^{−1} we obtain $\partial(T_\alpha - T_v)/\partial T_s = 3.8$. Thus, since GCM estimates of equilibrium tropical surface warming range from 1 to 4 K (ref. 21), this yields an increase in $(T_\alpha - T_v)$ of 3.8–11.4 K (accepting the one-dimensional model results would roughly double these numbers). This increase is large when compared with model predictions^{22,23} of the direct cooling effect in the Arctic vortex of doubling CO₂, which range from 2 K to 3.5 K. Shindell *et al.* showed that a 5 K cooling of the Arctic vortex led to a springtime destruction of ~50% of ozone within the Arctic vortex in the decade 2010–19, compared to ~15% depletion at present¹. Our results increase our confidence that at least this much cooling will occur.

Measurements of CH₄ concentration, and calculations of CH₄ sources and sinks, suggest that CH₄ concentrations may not grow much above current levels²⁴. If CH₄ concentrations were to resume growth at the rate observed during the 1980s, increasing CH₄ flux to the stratosphere might be expected to contribute another 1.3 p.p.p.v. of water vapour to the polar vortex mixing ratio by the year 2100. This would yield an additional increase of ~0.9 K in T_α .

Several studies have attempted to calculate trends based on observations of stratospheric water vapour. *In situ* frost-point hygrometer measurements yield a trend of 0.01–0.03 p.p.m.v. yr^{−1} for the years 1981–94 (ref. 25). Observations using a photofragment fluorescence hygrometer aboard the NASA ER-2 show no significant trend in stratospheric water vapour over the years 1993–97 (ref. 6), while observations from space show trends in water vapour mixing ratio of 0.04–0.10 p.p.m.v. yr^{−1} for the years 1991–96 (ref. 26). Since the last two trends are based on relatively short time series, they may have little to do with longterm climate variations. A trend of 0.04 p.p.m.v. yr^{−1} is equivalent to a tropopause temperature trend of 0.05 K yr^{−1} at the tropopause, so any connection between observed trends in stratospheric water vapour and trends in tropospheric climate will be difficult to detect, given the sparseness of tropical radiosonde measurements.

There are several sources of uncertainty in our calculations. The first is our continuing ignorance of the exact mechanism by which water vapour enters the stratosphere. If the cold-point tropopause temperature is actually controlled by convecting parcels reaching

that level, conservation of entropy by those parcels might lead to even larger increases in the cold-point tropopause temperature as surface temperature rises. Our results provide a strong motivation for the improvement of our understanding and observation of the tropical tropopause region and its sensitivity to changes in radiative and convective forcing. The second is the wide range of predicted changes in tropical surface temperature as greenhouse-gas concentrations increase. Last, a change in the stratospheric overturning circulation may affect polar chemistry in several ways: by changing the tropical tropopause temperature^{27,28}, by changing the temperature of persistence of the polar vortex¹, or by changing the average age of vortex air and thus the amount of water added by methane oxidation. All these effects need to be considered in assessing the effect of altered stratospheric dynamics. □

Received 15 March; accepted 18 October 1999.

- Shindell, D. T., Rind, D. & Loneragan, P. Increased polar stratospheric ozone losses and delayed eventual recovery owing to increasing greenhouse-gas concentrations. *Nature* **392**, 589–592 (1998).
- Austin, J., Butchart, N. & Shine, K. P. Possibility of an Arctic ozone hole in a doubled-CO₂ climate. *Nature* **360**, 221–225 (1992).
- Forster, P. M. & Shine, K. P. Stratospheric water vapour changes as a possible contributor to observed stratospheric cooling. *Geophys. Res. Lett.* (in the press).
- Chipperfield, M. P. & Pyle, J. A. Model sensitivity studies of Arctic ozone depletion. *J. Geophys. Res.* **103**, 28389–28403 (1998).
- Dessler, A. E. A reexamination of the “stratospheric fountain” hypothesis. *Geophys. Res. Lett.* **25**, 4165–4168 (1998).
- Hurst, D. F. *et al.* Closure of the total hydrogen budget of the northern extratropical lower stratosphere. *J. Geophys. Res.* **104**, 8191–8200 (1999).
- Anderson, J. G., Toohey, D. W. & Brune, W. H. Free-radicals within the Antarctic vortex—the role of CFCs in Antarctic ozone loss. *Science* **251**, 39–46 (1991).
- Müller, R. *et al.* Severe chemical ozone loss in the Arctic during the winter of 1995–96. *Nature* **389**, 709–712 (1997).
- Hofmann, D. J. & Oltmans, S. J. The effect of stratospheric water vapor on the heterogeneous reaction rate of ClONO₂ and H₂O for sulfuric acid aerosol. *Geophys. Res. Lett.* **19**, 2211–2214 (1992).
- Waibel, A. E. *et al.* Arctic ozone loss due to denitrification. *Science* **283**, 2064–2069 (1999).
- Carlsaw, K. S., Luo, B. & Peter, T. An analytic expression for the composition of aqueous HNO₃–H₂SO₄ stratospheric aerosols including gas phase removal of HNO₃. *Geophys. Res. Lett.* **22**, 1877–1880 (1995).
- Hanson, D. R. & Ravishankara, A. R. Reactive uptake of ClONO₂ onto sulfuric acid due to reaction with HCl and H₂O. *J. Phys. Chem.* **98**, 5728–5735 (1994).
- Kawa, S. R. *et al.* Activation of chlorine in sulfate aerosol as inferred from aircraft observations. *J. Geophys. Res.* **102**, 3921–3933 (1997).
- Schiller, C. *et al.* The partitioning of hydrogen species in the Arctic winter stratosphere: Implications for microphysical parameters. *J. Geophys. Res.* **101**, 14489–14493 (1996).
- Holton, J. R. *et al.* Stratosphere-troposphere exchange. *Rev. Geophys.* **33**, 403–439 (1995).
- Andrews, A. E. *et al.* Empirical age spectra for the lower tropical stratosphere from *in situ* observations of CO₂: Implications for stratospheric transport. *J. Geophys. Res.* (in the press).
- Hack, J. J., Kiehl, J. T. & Hurrell, J. W. The hydrologic and thermodynamic characteristics of the NCAR CCM3. *J. Clim.* **11**, 1179–1206 (1998).
- Emanuel, K. A. A scheme for representing cumulus convection in large-scale models. *J. Atmos. Sci.* **48**, 2313–2335 (1991).
- Hanson, M. D. & Suarez, M. J. An Efficient Thermal Radiation Parameterization for Use in General Circulation Models (NASA Tech. Memorandum 104606, Vol.3, NASA Goddard Space Flight Center, Greenbelt, Maryland, 1994).
- Hansen, J. *et al.* Forcing and chaos in interannual to decadal climate change. *J. Geophys. Res.* **102**, 25679–25720 (1997).
- Kattenberg, A. *et al.* in *Climate Change 1995* (eds Houghton, J. T. *et al.*) 285–358 (Cambridge Univ. Press, 1996).
- Rind, D., Suozzo, R., Balachandran, N. K. & Prather, M. J. Climate change and the middle atmosphere. Part I: The doubled CO₂ climate. *J. Atmos. Sci.* **47**, 475–494 (1990).
- Colman, R. A., Power, S. B., McAvaney, B. J. & Dahni, R. R. A non-flux corrected transient CO₂ experiment using the BMRC coupled atmosphere/ocean GCM. *Geophys. Res. Lett.* **22**, 3047–3050 (1995).
- Dlugokencky, E. J. *et al.* Continuing decline in the growth rate of the atmospheric CH₄ burden. *Nature* **393**, 447–450 (1998).
- Oltmans, S. J. & Hofmann, D. J. Increase in lower-stratospheric water vapour at a mid-latitude Northern Hemisphere site from 1981 to 1994. *Nature* **374**, 146–149 (1995).
- Nedoluha, G. E. *et al.* Increases in middle atmospheric water vapor as observed by the Halogen Occultation Experiment and the ground-based Water Vapor Millimeter-wave Spectrometer from 1991 to 1997. *J. Geophys. Res.* **103**, 3531–3543 (1998).
- Reid, G. C. & Gage, K. S. The tropical tropopause over the western Pacific: wave driving, convection, and the annual cycle. *J. Geophys. Res.* **101**, 21233–21241 (1996).
- Fells, S. B. Radiative-dynamical interactions in the middle atmosphere. *Adv. Geophys. A* **28**, 277–300 (1985).

Acknowledgements

We thank J. Kiehl for sharing output from the NCAR CCM3 GCM. This work was supported by NASA.

Correspondence and requests for materials should be addressed to D.B.K.-D. (e-mail: davidoff@huarp.harvard.edu). The updated model code from ref. 18 is available at <http://texmex.mit.edu/pub/emanuel/CONRAD>.

Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology

Pamela S. Soltis*, Douglas E. Soltis* & Mark W. Chase†

* School of Biological Sciences, Washington State University, Pullman, Washington 99164-4236, USA

† Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

Comparative biology requires a firm phylogenetic foundation to uncover and understand patterns of diversification and evaluate hypotheses of the processes responsible for these patterns. In the angiosperms, studies of diversification in floral form^{1,2}, stamen organization³, reproductive biology⁴, photosynthetic pathway⁵, nitrogen-fixing symbioses⁶ and life histories⁷ have relied on either explicit or implied phylogenetic trees. Furthermore, to understand the evolution of specific genes and gene families, evaluate the extent of conservation of plant genomes and make proper sense of the huge volume of molecular genetic data available for model organisms⁸ such as *Arabidopsis*, *Antirrhinum*, maize, rice and wheat, a phylogenetic perspective is necessary. Here we report the results of parsimony analyses of DNA sequences of the plastid genes *rbcl* and *atpB* and the nuclear 18S rDNA for 560 species of angiosperms and seven non-flowering seed plants and show a well-resolved and well-supported phylogenetic tree for the angiosperms for use in comparative biology.

Efforts to infer angiosperm phylogeny have greatly improved our understanding of the major lineages of flowering plants^{9–14}. However, despite these advances, the phylogenetic trees inferred from these studies are not completely congruent in the inter-relationships portrayed among the major lineages (although the alternatives are not strongly supported), and nearly all trees suffer from areas of poor resolution and/or weak support, usually due to low levels of divergence. In addition, molecular studies using hundreds of taxa present difficult analytical problems; the large number of taxa results in a huge number of possible trees¹⁵, increasing the amount of time needed to conduct a thorough analysis and decreasing the chances of finding the optimum tree(s). Phylogenetic analyses of 500 *rbcl* sequences did not find the most parsimonious trees⁹; further analyses of these data found slightly shorter trees^{16,17}. An analysis based on 18S rDNA¹¹ also probably failed to find the most parsimonious trees, in spite of extensive analyses. Thus, despite a decade of major improvements in our understanding of angiosperm phylogeny, the picture is not yet complete.

Table 1 Familiar taxa and the clades in which they are placed in the phylogenetic tree shown in Fig. 1

Familiar taxa/model organisms	Clade (Family or Order)	Larger group
<i>Nymphaea</i> (water lilies)	Nymphaeaceae	
<i>Magnolia</i>	Magnoliales	magnoliids
Orchidaceae (orchids)	Asparagales	monocots
Poaceae (grasses): <i>Triticum</i> , <i>Oryza</i> , <i>Zea</i>	Poales	monocots
Fabaceae (legumes)	Fabales	eurosid I
Rosaceae (rose family): <i>Rosa</i> , <i>Malus</i>	Rosales	eurosid I
<i>Arabidopsis</i> , <i>Brassica</i>	Brassicales	eurosid II
<i>Dianthus</i> (carnation), <i>Spinacia</i>	Caryophyllales	eudicots
<i>Cornus</i> (dogwoods)	Cornales	asterids
<i>Hydrangea</i>	Cornales	asterids
Lamiaceae (mints)	Lamiales	euasterid I
<i>Antirrhinum</i>	Lamiales	euasterid I
<i>Solanum</i> (potato, tomato)	Solanales	euasterid I
<i>Nicotiana</i> (tobacco)	Solanales	euasterid I
<i>Petunia</i>	Solanales	euasterid I
<i>Apium</i> (celery), <i>Daucus</i> (carrot)	Apiales	euasterid II
<i>Helianthus</i> (sunflower)	Asterales	euasterid II

The analytical issues involved in large-scale phylogenetic analyses have been discussed^{18,19}. Both simulations^{20,21} and empirical studies¹⁸ indicate that additional data can improve phylogenetic inferences from large data sets. Analyses of angiosperm relationships on the basis of gene sequences for *rbcl*, *atpB* and 18S rDNA for 190 angiosperm species and 3 outgroups showed increased resolution and internal support (as measured by bootstrap values), and faster run times when the data sets for these genes were combined rather than analysed separately. These studies indicated that additional data (that is, gene sequences) and taxa could improve inferences of angiosperm phylogeny.

Our analysis, which is based on parsimony analyses of 4,733 aligned nucleotides, provides a well-resolved and well-supported estimate of angiosperm phylogeny. Most of the major clades and some of the smaller ones recovered in this analysis were also found in previous studies^{9,11–13}. However, contrary to previous analyses based on data for one or two genes, all major clades and most of the spine of the tree receive jackknife (JK) support equal to or greater than 50%. Amborellaceae are the sister to all remaining angiosperms, consistent with results inferred from 18S rDNA¹¹ and *atpB*¹³ alone; the JK value for the clade of all angiosperms except *Amborella* is 65%. Nymphaeaceae (with a JK value of 100%) are then sister to all remaining angiosperms (JK 72%), followed by a clade of *Austrobaileya*, *Illicium* and *Schisandra* (JK for this clade of three genera 100%). This same branching order was found with stronger support in analyses based on five genes and nearly 9,000 base pairs (bp) of sequence data¹⁴ and is further corroborated by analyses of

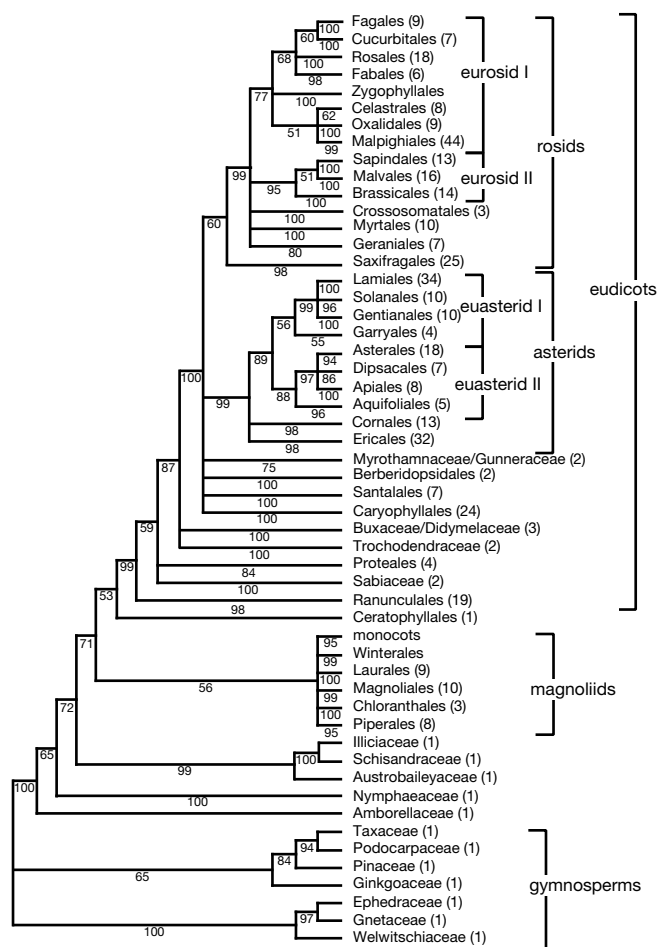


Figure 1 Summary of phylogenetic relationships for angiosperms inferred from analysis of *rbcl*, *atpB* and 18S rDNA sequences; the jackknife consensus tree (for groups receiving >50% support) is shown. The shortest trees found were 45,100 steps. The number of species in each clade is given in parentheses; not all 560 species occurred in a clade portrayed in this summary tree. Jackknife support is given below branches.

duplicate phytochrome genes²². The remaining angiosperms (JK 71%) form two major clades. One of these (JK 56%) consists of Chloranthaceae, Magnoliales, Laurales, Winterales and Piperales, all classified in subclass Magnoliidae²³, and the monocots. Each of these branches is strongly supported (JK $\geq 95\%$), but the relationships among these clades have JK values of less than 50%.

Within the monocots, *Acorus* is sister to a clade containing all other monocots (JK 99%). Alismatales are the next branching monocots and sister to a large clade (JK 99%) that comprises six main lineages: Petrosaviaceae, Dioscoreales, Pandanales, Liliales, Asparagales and commelinoids. Although all but Asparagales (JK 56%) and commelinoids (JK 68%) have JK percentages greater than 80%, relationships among clades of monocots are poorly resolved and/or weakly supported. The commelinoids, in turn, comprise Arecales (palms), Poales (including grasses), Commelinales and Zingiberales (gingers) as successive sister groups, although relationships among these four clades are not strongly supported. Only 102 out of the roughly 65,000 species of monocots²⁴ were included in this analysis; more extensive sampling of monocots and discussion of relationships are given in ref. 25.

The second major clade above the first three basal branches comprises *Ceratophyllum* as the weakly supported sister (JK 53%) to the eudicots, that is, all remaining angiosperms. The eudicots (JK 99%) comprise a series of successively branching orders: Ranunculales, Proteales, Trochodendrales, Gunneraceae/Myrothamnaceae, and a large clade of 'core eudicots' (JK 100%). This last clade contains the majority of all angiosperm species. Major clades within the core eudicots are Saxifragales as sister to the remaining rosids; Caryophyllales; and asterids, which comprise Cornales, Ericales, euasterids I and euasterids II. Model organisms in the rosids include *Arabidopsis*, *Brassica*, *Gossypium* and legumes. Nitrogen-fixing symbioses with nodulating bacteria arose within this clade and are confined to a portion of the eusoid I subclade. Although this 'nitrogen-fixing clade' has been found in previous analyses^{6,11,13}, this is the first study to our knowledge to provide strong support for it. Like the rosids, the asterid clade was recognized in previous studies^{9,11–13}, but without strong support. The model organisms *Antirrhinum*, *Nicotiana*, *Petunia*, *Solanum* and *Helianthus* are found in this large asterid clade. Familiar families found within each clade are given in Table 1. All three major clades of core eudicots identified here cut across subclass boundaries established in previous classifications^{23,24}. For the complete tree see Supplementary Information.

Although some minor portions of the tree are weakly supported, the first branches of angiosperm phylogeny, *Amborella*, Nymphaeales, *Austrobaileya*-*Illicium*-*Schisandra* (plus *Trimenia*¹⁴) now seem clear^{14,22}. The relationships among orders of magnoliids are only weakly supported, as is the position of the monocots. Within the eudicots, the branching order of the major clades is fairly clear, but levels of support could be increased by additional data. Finally, although the core eudicots form a clade (JK 100%), the relationships among rosids, Caryophyllales, asterids and a number of smaller clades lack substantial JK support. Additional data, both molecular and morphological, could increase support for these relationships.

Despite some areas of poor resolution and/or weak support, the general structure of angiosperm phylogeny is now clear and well supported. Using this framework, long-standing questions of plant diversification can be addressed, and recent discoveries can be placed in the proper historical context. For example, petals have apparently arisen multiple times within the angiosperms, from sepals in some cases to stamens in others^{1,26,27}, but the phylogenetic pattern of petal derivation is not clear. Using a phylogenetic tree such as ours, it would be possible to assess the homology of 'petals' throughout the angiosperms and then to search for genetic and developmental mechanisms of 'petal' formation²⁸. Other aspects of floral evolution, such as the origin, maintenance and diversification

of synorganized flowers¹ in the core eudicots and patterns of stamen organization³, can also be addressed. Inferences of diversification of other angiosperm characteristics, such as morphology, physiology, ecology and genome structure, are now possible on a much more robust basis than previously possible. Although large analyses of the angiosperms have been reported^{9,11–13}, this study provides both greater resolution and stronger support for the relationships inferred. The use of broad taxon sampling and many characters limits spurious results due to homoplasy and unequal rates of evolution among taxa and provides stronger support for relationships because historical information from all three genes is combined. Our tree, which employs three times the quantity of data used to construct most other trees, provides a substantially more secure foundation for inferring the evolution of key features in the angiosperms. □

Methods

Selection of species was designed to sample all major groups of angiosperms, using recent classifications^{23,24} and phylogenetic studies^{11,13,16} as guides. Sequences of *rbcl*, *atpB*, and 18S rDNA were amplified by polymerase chain reaction (PCR) and sequenced using an ABI 377 automated DNA sequencer. Details of sampling, DNA extraction, PCR amplification, sequencing and alignment will be published elsewhere (D.E.S. *et al.*, manuscript in preparation). Voucher information, GenBank numbers and the aligned data matrix are available at <http://www.wsu.edu:8080/~soltislab/>, <http://www.kewgardens.org> and <http://www.ucjeps.berkeley.edu/bryolab/greenplantpage.html>. Heuristic parsimony analyses were conducted using PAUP* 4.0 (ref. 29) and the ratchet (K. Nixon) with NONA (P. Goloboff). Parsimony JK analyses³⁰ (1,200 replicates, each with ten random-entry order replicates and branch swapping) were used to measure support for the topology. The gymnosperms *Ephedra*, *Gnetum*, *Welwitschia*, *Ginkgo*, *Pinus*, *Podocarpus* and *Taxus* were specified as the outgroup.

Received 6 September; accepted 19 October 1999.

- Endress, P. K. Floral structure and evolution of primitive angiosperms: recent advances. *Plant Syst. Evol.* **192**, 79–97 (1994).
- Donoghue, M. J., Ree, R. H. & Baum, D. A. Phylogeny and the evolution of flower symmetry in the Asteridae. *Trends Plant Science* **3**, 311–317 (1998).
- Ronse Decraene, L.-P. & Smets, E. The distribution and systematic relevance of the androecial characters oligomery and polymery in the Magnoliophytina. *Nord. J. Bot.* **7**, 239–253 (1987).
- Weller, S. G., Donoghue, M. J. & Charlesworth, D. In *Experimental and Molecular Approaches to Plant Biosystematics* (eds Hoch, P. C. & Stephenson, A. G.) 355–382 (Missouri Botanical Garden, St. Louis, 1995).
- Kellogg, E. A. In *The Biology of C4 Photosynthesis* (eds Sage, R. F. & Monson, R. K.) 411–444 (Academic Press, San Diego, 1998).
- Soltis, D. E. *et al.* Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl Acad. Sci. USA* **92**, 2647–2651 (1995).
- Dodd, M. E., Silvertown, J. & Chase, M. W. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* **53**, 732–744 (1999).
- Gale, M. D. & Devos, K. M. Plant comparative genetics after 10 years. *Science* **282**, 656–658 (1998).
- Chase, M. W. *et al.* Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcl*. *Ann. Missouri Bot. Gard.* **80**, 528–580 (1993).
- Doyle, J. A., Donoghue, M. J. & Zimmer, E. A. Integration of morphological and ribosomal RNA data on the origin of angiosperms. *Ann. Missouri Bot. Gard.* **81**, 419–450 (1994).
- Soltis, D. E. *et al.* Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* **84**, 1–49 (1997).
- Nandi, W. I., Chase, M. W. & Endress, P. K. A combined cladistic analysis of angiosperms using *rbcl* and non-molecular data sets. *Ann. Missouri Bot. Gard.* **85**, 137–212 (1998).
- Savolainen, V. *et al.* Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcl* gene sequences. *Syst. Biol.* (in the press).
- Qiu, Y.-L. *et al.* The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* **402**, 404–407 (1999).
- Felsenstein, J. The number of evolutionary trees. *Syst. Zool.* **27**, 27–33 (1978).
- Rice, K. A., Donoghue, M. J. & Olmstead, R. G. Analyzing large data sets: *rbcl* 500 revisited. *Syst. Biol.* **46**, 554–563 (1997).
- Nixon, K. C., Davis, J. L. & Goloboff, P. A. Search strategies for large datasets: an example using *rbcl*. *Am. J. Bot.* **85**, 148 (1998).
- Soltis, D. E. *et al.* Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. *Syst. Biol.* **47**, 32–42 (1998).
- Chase, M. W. & Cox, A. V. Gene sequences, collaboration and analysis of large data sets. *Australian Syst. bot.* **11**, 215–229 (1998).
- Hillis, D. M. Inferring complex phylogenies. *Nature* **383**, 130 (1996).
- Graybeal, A. Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* **47**, 9–17 (1998).
- Mathews, S. & Donoghue, M. J. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* **286**, 947–950 (1999).
- Cronquist, A. *An Integrated System of Classification of Flowering Plants* (Columbia University Press, New York, 1981).
- Takhtajan, A. *Diversity and Classification of Flowering Plants* (Columbia University Press, New York, 1997).
- Chase, M. W. *et al.* Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification, in *Proceedings of Monocots II: The Second International Symposium*

- on the Comparative Biology of the Monocotyledons, Sydney, Australia, (eds Wilson, K. & Morrison, D.) (CSIRO Press, Sydney, in the press).
26. Cronquist, A. *The Evolution and Classification of Flowering Plants* (The New York Botanical Garden, New York, 1988).
 27. Takhtajan, A. *Evolutionary Trends in Flowering Plants* (Columbia University Press, New York, 1991).
 28. Kramer, E. M., Dorit, R. L. & Irish, V. F. Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* **149**, 765–783 (1998).
 29. Swofford, D. L. *Phylogenetic Analysis Using Parsimony** (PAUP*), version 4.0. (Sinauer Associates, Sunderland, MA, 1998).
 30. Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D. & Kluge, A. G. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12**, 99–124 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

This work was supported in part by the NSF.

Correspondence and requests for material should be addressed to P.S.S. (e-mail: psoltis@wsu.edu).

The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes

Yin-Long Qiu*, Jungho Lee*, Fabiana Bernasconi-Quadroni*, Douglas E. Soltis†, Pamela S. Soltis†, Michael Zanis†, Elizabeth A. Zimmer‡, Zhiduan Chen*, Vincent Savolainen||, Mark W. Chase||

* Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland

† School of Biological Sciences, Washington State University, Pullman, Washington 99164-4236, USA

‡ Laboratory of Molecular Systematics, Smithsonian Institution, Washington DC 20560, USA

|| Jodrell Laboratory Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

Angiosperms have dominated the Earth's vegetation since the mid-Cretaceous (90 million years ago)¹, providing much of our food, fibre, medicine and timber, yet their origin and early evolution have remained enigmatic for over a century^{2–8}. One part of the enigma lies in the difficulty of identifying the earliest angiosperms; the other involves the uncertainty regarding the sister group of angiosperms among extant and fossil gymnosperms. Here we report a phylogenetic analysis of DNA sequences of five mitochondrial, plastid and nuclear genes (total aligned length 8,733 base pairs), from all basal angiosperm and gymnosperm lineages (105 species, 103 genera and 63 families). Our study demonstrates that *Amborella*, Nymphaeales and Illiciales-*Trimeniaceae-Austrobaileya* represent the first stage of angiosperm evolution, with *Amborella* being sister to all other angiosperms. We also show that Gnetales are related to the conifers and are not sister to the angiosperms, thus refuting the Anthophyte Hypothesis¹. These results have far-reaching implications for our understanding of diversification, adaptation, genome evolution and development of the angiosperms.

Difficulty in identifying the earliest angiosperms is the result of three problems that characterize diversification of most major clades. First, the great divergence between gymnosperms and angiosperms makes assessment of character homology difficult and thus renders the otherwise powerful outgroup-approach prob-

lematic in morphological cladistic analyses^{1,9}. Second, extinction, which is partly responsible for this divergence, has almost certainly occurred in both groups^{1,10–16}, and highlights the need of extensive taxon sampling when relying on the living diversity. Last, the fossil evidence indicates that the early angiosperms went through an explosive radiation^{1,10–16}, which to resolve requires the sampling of a large number of characters. Previous molecular analyses have had some success in resolving relationships among basal angiosperms^{17–21}; however, their results are only weakly supported, and worse, are often contradictory because of evolutionary rate heterogeneity among lineages of the particular gene used, weak phylogenetic signal in single genes, and insufficient taxon sampling. From both theoretical and empirical studies, it is becoming increasingly clear that to address such a difficult issue as basal angiosperm phylogeny, extensive sampling in both dimensions of taxa and characters (genes) is necessary^{22–24}.

We obtained sequences of five genes from all three plant genomes: mitochondrial *atp1* and *matR*, plastid *atpB* and *rbcL* and nuclear 18S rDNA. They encode products involved in energy metabolism, carbohydrate synthesis and information processing. Thus, our character sampling strategy of taking multiple genes of different functions from all three genomes is designed to reduce homoplasy generated by gene-, function- and genome-specific molecular evolutionary phenomena such as rate heterogeneity, GC content bias, RNA editing and protein structural constraints^{25,26}. To optimize the performance of phylogenetic methods in analysing complex diversification patterns in early angiosperms^{22,23}, we included 97 species, 95 genera and 55 families of basal angiosperms, essentially sampling all living families^{5,8,11,20,21}. Eight gymnosperms from eight families were used as outgroups. The DNA sequences were analysed with parsimony methods; bootstrap (BS) and jackknife (JK) analyses were conducted to measure stability of phylogenetic patterns.

The same single most parsimonious tree was found in each of 1,000 random taxon-addition replicates in the analysis (Fig. 1). *Amborella*, a shrub of the monotypic New Caledonian family Amborellaceae, is sister to all other angiosperms, which are strongly (90% BS and 92% JK) supported as a monophyletic group. The next diverging lineage corresponds to Nymphaeales, the water lilies; its sister clade of the remaining angiosperms receives 98% BS and 99% JK support. The third clade consists of two small Australasian families, Austrobaileyaaceae and Trimeniaceae, and two small eastern Asia-eastern North America disjunct families, Illiciaceae and Schisandraceae (Illiciales). All remaining angiosperms (euangiosperms) make up a strongly supported large clade (97% BS and 99% JK). The relationships among lineages within euangiosperms are resolved in the shortest tree but generally receive less than 50% BS support. All major lineages, however, are strongly supported; these agree with previous classifications^{5,8,11} and results of cladistic analyses of morphological and molecular data^{9,20,21,27}. Among gymnosperms, two gnetalean genera, *Gnetum* and *Welwitschia*, are not sister to angiosperms as suggested by the Anthophyte Hypothesis¹, but fall close to the conifers.

We observed one INDEL (insertion/deletion) in *matR* that supports the basal position of *Amborella*, Nymphaeales and Illiciales-*Trimeniaceae-Austrobaileya* (ANITA) in angiosperms: an 18-base-pair (bp) deletion in all euangiosperms but not in ANITA or gymnosperms, some of which have 6–15-bp deletions (Fig. 2). Although we cannot rule out the possibility that the sequence in the INDEL region of ANITA and gymnosperms results from independent insertions, two lines of evidence suggest that this scenario is unlikely. The sequences are found in all three ANITA lineages and all four gymnosperm lineages. Furthermore, there are identical or similar codons shared by ANITA and gymnosperms in the INDEL.

Reconstruction of deep phylogenies using DNA sequences has been plagued by problems caused by rate heterogeneity, weak phylogenetic signal in single genes, insufficient taxon sampling,

§ Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, 100093, China.

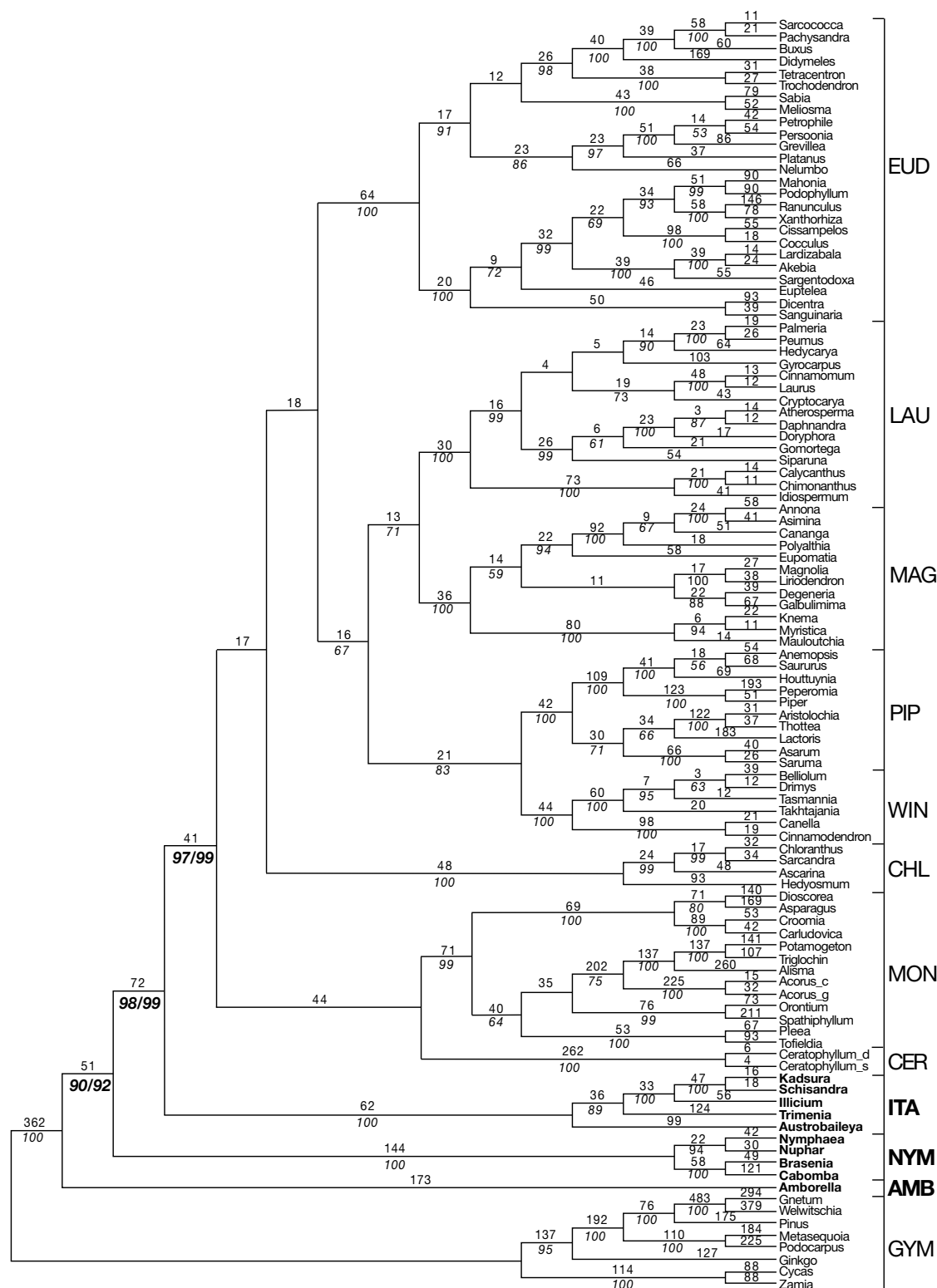


Figure 1 The single most parsimonious tree found in the five-gene DNA sequence analysis (tree length, 13,240 steps; consistency index, 0.413; retention index, 0.604). Numbers above branches are branch lengths (ACCTRAN optimization); those below in italics are bootstrap values (only those above 50% are shown; for branches related to ANITA (bold type), numbers below branches before the slash are bootstrap values and

those after are jackknife values). GYM, gymnosperms; AMB, *Amborella*; NYM, Nymphaeales; ITA, Illiciales, Trimeniaceae and *Austrobaileya*; CER, *Ceratophyllum*; MON, monocots; CHL, Chloranthaceae; WIN, Winterales; PIP, Piperales; MAG, Magnoliales; LAU, Laurales; EUD, eudicots; *Acorus_c*, *A. calamus*; *Acorus_g*, *A. gramineus*; *Ceratophyllum_d*, *C. demersum*; *Ceratophyllum_s*, *C. submersum*.

explosive radiation, extinction and protein structural constraints^{25,26}. In retrospect, our earlier studies using single genes suffered from some of these problems when *Ceratophyllum* was found to be sister to all other angiosperms^{20,21}. The same concern

could still be raised about our results presented here; however, the use of five genes with different functions from all three genomes and the sampling of almost all basal angiosperm families is likely to have reduced considerably the effect of these problems. Five lines of

Arabidopsis	AAT	AAT	AAT	---	---	---	---	---	---	TGG	---	GCC
Oenothera	...	...	...	---	---	---	---	---	---	T.	---	---
Vicia	...	...	...	---	---	---	---	---	---	T.	---	---
Solanum	...	...	...	---	---	---	---	---	---	T.	---	---
Triticum	...	...	...	---	---	---	---	---	---	T.	---	---
Sarcococca	...	...	...	---	---	---	---	---	---	---	---	---
Pachysandra	...	...	...	---	---	---	---	---	---	---	---	---
Buxus	...	...	...	---	---	---	---	---	---	---	---	---
Didymeleis	...	...	...	---	---	---	---	---	---	G.	---	---
Tetracentron	...	...	...	---	---	---	---	---	---	---	---	---
Trochodendron	...	...	...	---	---	---	---	---	---	---	---	---
Sabia	...	...	...	---	---	---	---	---	---	---	---	---
Petrophile	...	...	...	---	---	---	---	---	---	---	---	---
Persoonia	...	...	...	---	---	---	---	---	---	---	---	---
Grevillea	...	...	...	---	---	---	---	---	---	---	---	---
Platanus	...	...	...	---	---	---	---	---	---	---	---	---
Nelumbo	...	...	...	---	---	---	---	---	---	---	---	---
Mahonia	...	...	...	---	---	---	---	---	---	---	---	---
Podophyllum	...	...	...	---	---	---	---	---	---	---	---	---
Ranunculus	...	...	...	---	---	---	---	---	---	T.G	---	---
Xanthorhiza	...	...	...	---	---	---	---	---	---	T.G	---	---
Cissampelos	T.	...	...	---	---	---	---	---	---	---	---	---
Cocculus	...	...	...	---	---	---	---	---	---	---	---	---
Lardizabala	...	...	...	---	---	---	---	---	---	---	---	---
Akebia	...	...	...	---	---	---	---	---	---	T.A	---	---
Sargentodoxa	...	...	...	---	---	---	---	---	---	T.A	---	---
Euptelea	...	...	...	---	---	---	---	---	---	T.A	---	---
Dicentra	...	...	...	---	---	---	---	---	---	---	---	---
Sanguinaria	...	...	...	---	---	---	---	---	---	---	---	---
Palmeria	...	...	...	---	---	---	---	---	---	T.	---	---
Peumus	...	...	...	---	---	---	---	---	---	---	---	---
Hedycarya	...	...	...	---	---	---	---	---	---	---	---	---
Gyrocarpus	...	...	...	---	---	---	---	---	---	---	---	---
Cinnamomum	T.	...	...	---	---	---	---	---	---	---	---	---
Laurus	...	...	...	---	---	---	---	---	---	---	---	---
Cryptocarya	...	...	...	---	---	---	---	---	---	G.	---	---
Atroposperma	...	...	...	---	---	---	---	---	---	---	---	---
Daphnandra	...	...	...	---	---	---	---	---	---	---	---	---
Doryphora	...	...	...	---	---	---	---	---	---	---	---	---
Siparuna	...	...	...	---	---	---	---	---	---	---	---	---
Calycanthus	...	T.	...	---	---	---	---	---	---	---	---	---
Chimonanthus	...	T.	...	---	---	---	---	---	---	---	---	---
Idiospermum	...	A	C.	---	---	---	---	---	---	---	---	---
Annona	C.	G.	...	---	---	---	---	---	---	T.C	---	---
Asimina	...	G.	...	---	---	---	---	---	---	T.A	---	---
Cananga	...	G.	...	---	---	---	---	---	---	T.A	---	---
Polyalthia	...	G.	...	---	---	---	---	---	---	T.A	---	---
Eupomatia	...	G.	...	---	---	---	---	---	---	T.A	---	---
Magnolia	...	...	...	---	---	---	---	---	---	---	---	---
Liriodendron	...	...	...	---	---	---	---	---	---	---	---	---
Degeneria	...	...	A	---	---	---	---	---	---	T.	---	---
Galbulimima	...	G.	...	---	---	---	---	---	---	T.A	---	---
Knema	...	G.	...	---	---	---	---	---	---	---	---	T.
Myristica	...	G.	...	---	---	---	---	---	---	T.A	---	---
Mauloutchia	...	G.	...	---	---	---	---	---	---	T.T	---	---
Anemopsis	...	...	...	---	---	---	---	---	---	---	---	A.
Saururus	T.	T.G	T.	ATA	TAT	ATA	TAT	ACA	TAT	T.	---	A.
Houttuynia	...	T.G	T.	ATA	TAT	ATA	TAT	ACA	TAT	---	---	A.
Peperomia	...	...	G.	---	---	---	---	---	A.	---	---	A.
Piper	...	...	...	---	---	---	---	---	---	---	---	---
Aristolochia	...	...	...	---	---	---	---	---	---	---	---	---
Thottea	...	...	...	---	---	---	---	---	---	---	---	---
Lactoris	C.	...	...	---	---	---	---	---	---	---	---	---
Asarum	...	...	...	---	---	---	---	---	---	---	---	---
Saruma	...	...	...	---	---	---	---	---	---	---	---	---
Belicolum	T.	C.	A	---	---	---	---	---	---	T.G	---	---
Drimys	...	...	...	---	---	---	---	---	---	---	---	---
Tasmannia	...	...	...	---	---	---	---	---	---	---	---	---
Canella	C.	...	...	---	---	---	---	---	---	---	---	---
Cinnamodendron	C.	...	...	---	---	---	---	---	---	---	---	---
Chloranthus	C.	...	...	---	---	---	---	---	---	---	---	---
Sarcandra	C.	...	...	---	---	---	---	---	---	---	---	---
Ascarina	C.	...	...	---	---	---	---	---	---	---	---	---
Hedyosmum	C.	...	...	---	---	---	---	---	---	---	---	---
Dioscorea	...	G.	...	---	---	---	---	---	---	T.	---	---
Asparagus	...	G.	...	---	---	---	---	---	---	T.	---	---
Croomia	...	G.	...	---	---	---	---	---	---	T.	---	---
Carludovica	...	G.	...	---	---	---	---	---	---	T.	---	---
Potamogeton	G.G	G.	...	---	---	---	---	---	---	T.	---	---
Triglochin	G.G	G.	...	---	---	---	---	---	---	T.	---	---
Alisma	G.G	G.	...	---	---	---	---	---	---	T.	---	---
Orontium	...	G.	...	---	---	---	---	---	---	T.	---	---
Spathiphyllum	...	A	---	---	---	---	---	---	---	T.	---	---
Pilea	...	G.	...	---	---	---	---	---	---	T.	---	---
Tofieldia	...	G	...	---	---	---	---	---	---	T.	---	---
Ceratophyllum_d	T.	C.	A	---	---	---	---	---	---	G.	---	---
Ceratophyllum_s	T.	C.	A	---	---	---	---	---	---	G.	---	---
Kadsura	T.	T.G	T.	ATA	TAT	ATA	TAT	ACA	TAT	G.	---	---
Schisandra	T.	T.G	T.	ATA	TAT	ATA	TAT	ACA	TAT	G.	---	---
Illicium	T.	T.A	T.	ACA	TAT	ATA	---	---	---	G.	---	---
Trimenia	T.	T.A	T.	ATA	TAT	ATA	TAT	---	---	G.	---	---
Austrobaileya	G.	T.A	T.	ATA	TAT	AGA	TAT	---	---	G.	---	---
Nymphaea	...	...	...	AAT	---	---	---	---	---	G.	---	---
Nuphar	...	...	...	AAT	---	---	---	---	---	G.	---	---
Brasenia	...	C.G	...	TAT	---	---	---	---	---	GA.	---	---
Cabomba	...	C.G	CT.	ACA	GTA	ACG	CAA	CCA	GTA	A.T	---	---
Amborella	...	...	TCA	AAT	---	ACG	---	CCA	---	G.	---	---
Gnetum	T.C	T.	...	ACT	ACC	CGG	---	---	---	G.	GGC	ATA
Welwitschia	T.	...	...	CC	CAG	CAT	AAT	---	---	AGT	G.	GAC
Pinus	...	...	...	AAT	AAT	GGG	---	---	---	GTA	GC.	TTC
Ginkgo	...	G.	...	GAT	AAT	AGG	---	---	---	GTA	GC.	TTC
Cycas	...	G.	G	GAG	AAC	AGG	---	---	---	GTA	GC.	TTC
Zamia	...	G.	G.G	GAT	AAC	AGG	---	---	---	GTA	GC.	TTC

Figure 2 The portion of the aligned matrix from mitochondrial *matR* showing the INDEL that distinguishes euangiosperms (top block) from ANITA (middle block) and gymnosperms (bottom block). Dots indicate nucleotides identical to the top sequence, and dashes indicate gaps. Sequences of *Arabidopsis*, *Oenothera*, *Vicia*, *Solanum* and *Triticum* (from GenBank; not used in the phylogenetic analysis) are included here to show the INDEL status in derived eudicots and monocots. The codon grouping shown is the correct reading frame. Abbreviations as in Fig. 1.

independent evidence support our identification of ANITA as the earliest angiosperms from the DNA sequence analysis. First, we carried out five single-gene analyses. Three of them, *atp1*, *matR* and *atpB*, placed ANITA at the base of angiosperms, with the exception that in the *atp1* analysis, *Amborella* fell into eudicots because of its divergent sequence. This example, together with that of *Ceratophyllum* in *rbcL* analyses^{20,21}, exposes the weakness of single-gene analysis, the very reason we conducted this multigene analysis. The 18S rDNA analysis yielded a polytomy among basal angiosperms (but an earlier study of 18S rDNA sampling 223 seed plants identified ANITA as the earliest angiosperms¹⁹), whereas analysis of *rbcL* still rooted the angiosperm tree at *Ceratophyllum*, consistent with earlier results^{20,21} (removal of *Ceratophyllum*, however, changed the root to ANITA). We also analysed amino-acid sequences of the four protein-coding genes and still found ANITA at the base of angiosperms.

Second, the INDEL in *matR*, being a non-point-mutation character and thus less prone to convergence than nucleotide substitutions, clearly separates ANITA from euangiosperms and places them closer to the gymnosperms. Third, a study using duplicated phytochrome genes to root the angiosperm phylogeny and two analyses of multiple genes with different taxon sampling have independently corroborated that ANITA are the earliest angiosperms^{27,28,31}. Fourth, all ANITA members (except Illiciaceae) share one morphological feature: carpel closure at anthesis through occlusion of the inner space by secretion. This feature is rare in euangiosperms and is probably a primitive condition in the earliest angiosperms²⁹. Last, fossil evidence, although awaiting further discovery, is beginning to show a pattern consistent with our topology. Many floral structures from the Early Cretaceous show similarities to those of ANITA^{1,12,13}. Taken together, these lines of evidence provide strong and clear support for this deep split of angiosperms.

Identifying the earliest angiosperms solves one of the two critical pieces of the angiosperm origin mystery², and will contribute to solving the other piece: determining the sister group of angiosperms among extant and fossil gymnosperms. Our analyses indicate that the prevalent view of the past two decades that Gnetales are sister to angiosperms is incorrect. These results will also help elucidate the evolution of several features that contributed to the rise and ultimate dominance of angiosperms in modern terrestrial ecosystems, for example, floral development, carpel closure, insect pollination and double fertilization. Finally, these findings will allow selection of appropriate model organisms to investigate critical changes in genomic evolution and developmental programmes, which have led to the overwhelming diversity we see today.

This study also has important implications for future work on reconstructing ancient phylogenies. DNA sequence analyses, although proven to be powerful in reconstructing recent evolutionary histories, have often failed to resolve difficult, long-standing controversies in organismal diversification patterns²⁵. Although many problems may be responsible for this failure, one obvious aspect of experimental design has been paradoxically overlooked: collection of enough data to reduce sampling error. Several studies experimenting with extensive sampling of either taxa or characters have been published^{17,19,20,26}, but few with both (see ref. 27). Our analysis demonstrates that extensive sampling of both taxa and characters is important for resolving difficult phylogenetic issues. □

Methods

Gene sequencing

Total cellular DNA was extracted from fresh or silica-gel-dried leaves using the standard CTAB method (see ref. 21). The genes were amplified by conventional PCR and sequenced using an ABI-PRISM 377 DNA sequencer (PE Applied Biosystems) according to manufacturer's protocols with modifications. The five genes were aligned individually, using XPILEUP from the GCG package (Genetics Computer Group, Inc.) with different settings for gap-creation penalty and gap-extension penalty. Minor manual adjustments were made after computer alignment.

Phylogenetic analyses

For the taxa analysed, all 105 species had *rbcl* sequences, and 93, 88, 98 and 101 species had *atpB*, 18S rDNA, *matR* and *atp1* sequences, respectively (missing data for critical taxa: *Kadsura*: 18S rDNA, *Trimeria*: *atp1*, *Cycas* and *Zamia*: *atpB*, and *Metasequoia* and *Podocarpus*: *matR*). Each taxon had data for at least three out of the five genes. Parsimony (equal weighting) analyses were carried out using PAUP*4.0b2 (ref. 30). To search for islands of shortest trees, a heuristic search was conducted using 1,000 random taxon-addition replicates, one tree held at each step during stepwise addition, TBR branch swapping, steepest descent option in effect. MulTrees option in effect and no upper limit of MaxTrees. Both bootstrap and jackknife (50% character deletion) analyses were conducted using 1,000 resampling replicates and the same tree search procedure as described above except with simple taxon addition. The data matrix is available as Supplementary Information at <http://www.nature.com>.

All *atp1* and *matR*, and some *atpB*, *rbcl* and 18S rDNA sequences were generated in this study, deposited in GenBank under accession numbers AF197576–AF197815; remaining sequences were from GenBank and ref. 27.

- Crane, P. R., Friis, E. M. & Pedersen, K. R. The origin and early diversification of angiosperms. *Nature* **374**, 27–33 (1995).
- Darwin, C. in *More Letters of Charles Darwin: A Record of His Work in a Series of Hitherto Unpublished Letters* Vol. 2 (eds Darwin, F. & Seward, A. C.) 20–22, 26–27 (John Murray, London, 1903).
- Arber, E. A. N. & Parkin, J. On the origin of angiosperms. *Bot. J. Linnean Soc.* **38**, 29–80 (1907).
- von Wettstein, R. R. *Handbuch der Systematischen Botanik. II. Band* (Franz Deuticke, Wien, 1907).
- Takhtajan, A. *Flowering Plants: Origin and Dispersal* (Oliver and Boyd, Edinburgh, 1969).
- Doyle, J. A. Origin of angiosperms. *Annu. Rev. Ecol. Syst.* **9**, 365–392 (1978).
- Endress, P. K. Reproductive structures and phylogenetic significance of extant primitive angiosperms. *Pl. Syst. Evol.* **152**, 1–28 (1986).
- Cronquist, A. *The Evolution and Classification of Flowering Plants* 2nd edn (The New York Botanical Garden, New York, 1988).
- Donoghue, M. J. & Doyle, J. A. in *Evolution, Systematics, and Fossil History of the Hamamelidae* Vol. 1 (eds Crane, P. R. & Blackmore, S.) 17–45 (Clarendon, Oxford, 1989).
- Doyle, J. A. Cretaceous angiosperm pollen of the Atlantic Coastal Plain and its evolutionary significance. *J. Arnold Arb.* **50**, 1–35 (1969).
- Walker, J. W. & Walker, A. G. Ultrastructure of lower Cretaceous angiosperm pollen and the origin and early evolution of flowering plants. *Ann. Missouri Bot. Gard.* **71**, 464–521 (1984).
- Friis, E. M., Pedersen, K. R. & Crane, P. R. Angiosperm floral structures from the Early Cretaceous of Portugal. *Pl. Syst. Evol. (Suppl.)* **8**, 31–49 (1994).
- Friis, E. M., Pedersen, K. R. & Crane, P. R. Early angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. *Ann. Missouri Bot. Gard.* **86**, 259–296 (1999).
- Walker, J. W., Brenner, G. J. & Walker, A. G. Winteraceous pollen in the lower Cretaceous of Israel: early evidence of a magnolioid angiosperm family. *Science* **220**, 1273–1275 (1983).
- Taylor, D. W. & Hickey, L. J. An Aptian plant with attached leaves and flowers: implications for angiosperm origin. *Science* **247**, 702–704 (1990).
- Sun, G., Dilcher, D. L., Zheng, S. & Zhou, Z. In search of the first flower: a Jurassic angiosperm, *Archaeofructus*, from Northeast China. *Science* **282**, 1692–1695 (1998).
- Martin, P. G. & Dowd, J. M. Studies of angiosperm phylogeny using protein sequences. *Ann. Missouri Bot. Gard.* **78**, 296–337 (1991).
- Hamby, R. K. & Zimmer, E. A. in *Molecular Systematics of Plants* (eds Soltis, P. S., Soltis, D. E. & Doyle, J. J.) 50–91 (Chapman and Hall, New York, 1992).
- Soltis, D. E. *et al.* Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* **84**, 1–49 (1997).
- Chase, M. W. *et al.* Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcl*. *Ann. Missouri Bot. Gard.* **80**, 528–580 (1993).
- Qiu, Y.-L., Chase, M. W., Les, D. H. & Parks, C. R. Molecular phylogenetics of the Magnoliidae: cladistic analyses of nucleotide sequences of the plastid gene *rbcl*. *Ann. Missouri Bot. Gard.* **80**, 587–606 (1993).
- Hillis, D. M. Inferring complex phylogenies. *Nature* **383**, 130–131 (1996).
- Graybeal, A. Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* **47**, 9–17 (1998).
- Soltis, D. E. *et al.* Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. *Syst. Biol.* **47**, 32–42 (1998).
- Qiu, Y.-L. & Palmer, J. D. Phylogeny of early land plants: insights from genes and genomes. *Trends Plant Sci.* **4**, 26–30 (1999).
- Naylor, G. J. P. & Brown, W. M. Structural biology and phylogenetic estimation. *Nature* **388**, 527–528 (1997).
- Soltis, P. S., Soltis, D. E. & Chase, M. W. Angiosperm phylogeny inferred from multiple genes as a research tool for comparative biology. *Nature* **402**, 402–404 (1999).
- Mathews, S. & Donoghue, M. J. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* **286**, 947–950 (1999).
- Endress, P. K. & Igersheim, A. Gynoecium diversity and systematics of the Laurales. *Bot. J. Linnean Soc.* **125**, 93–168 (1997).
- Swofford, D. L. PAUP*4.0b2: Phylogenetic Analysis Using Parsimony. (Sinauer, Sunderland, Massachusetts, 1998).
- Parkinson, C. L., Adams, K. L. & Palmer, J. D. Multigene analyses identify the three earliest lineages of extant flowering plants. *Curr. Biol.* (in the press).

Acknowledgements

We thank C. D. K. Cook, M. E. Endress, P. K. Endress, E. M. Friis, O. Nandi and R. Rutishauser for critical reading of the manuscript, R. Collett, A. Floyd, B. Hall and S. S. Renner for plant material, and the Swiss NF and US NSF for financial support.

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Correspondence and requests for materials should be addressed to Y.-L.Q. (e-mail: yqiu@systbot.unizh.ch; or after February 2000, yqiu@bio.umass.edu).

Biodiversity of plankton by species oscillations and chaos

Jef Huisman*†‡ & Franz J. Weissing§

* Biological Sciences, Stanford University, Stanford, California 94305-5020, USA

† Center for Estuarine and Marine Ecology, CEMO-NIOO, PO Box 140, 4400 AC Yerseke, The Netherlands

§ Department of Genetics, University of Groningen, PO Box 14, 9750 AA Haren, The Netherlands

Biodiversity has both fascinated and puzzled biologists¹. In aquatic ecosystems, the biodiversity puzzle is particularly troublesome, and known as the 'paradox of the plankton'². Competition theory predicts that, at equilibrium, the number of coexisting species cannot exceed the number of limiting resources^{3–6}. For phytoplankton, only a few resources are potentially limiting: nitrogen, phosphorus, silicon, iron, light, inorganic carbon, and sometimes a few trace metals or vitamins. However, in natural waters dozens of phytoplankton species coexist². Here we offer a solution to the plankton paradox. First, we show that resource competition models^{6–10} can generate oscillations and chaos when species compete for three or more resources. Second, we show that these oscillations and chaotic fluctuations in species abundances allow the coexistence of many species on a handful of resources. This model of planktonic biodiversity may be broadly applicable to the biodiversity of many ecosystems.

We consider a well-known resource competition model^{6–10} that has been tested and verified extensively using competition experiments with phytoplankton species^{8,11–16}. Consider n species and k resources. Let N_i denote the population abundance of species i , and let R_j denote the availability of resource j . The dynamics of the species depend on the availabilities of the resources. The resource availabilities, in turn, depend on the rates of resource supply and the amount of resources consumed by the phytoplankton species. This gives the following model^{6–9}:

$$\frac{dN_i}{dt} = N_i(\mu_i(R_1, \dots, R_k) - m_i) \quad i = 1, \dots, n \quad (1)$$

$$\frac{dR_j}{dt} = D(S_j - R_j) - \sum_{i=1}^n c_{ji}\mu_i(R_1, \dots, R_k)N_i \quad j = 1, \dots, k \quad (2)$$

Here $\mu_i(R_1, \dots, R_k)$ is the specific growth rate of species i as a function of the resource availabilities; m_i is the specific mortality rate of species i ; D is the system's turnover rate; S_j is the supply concentration of resource j ; and c_{ji} is the content of resource j in species i . We assume that the specific growth rates follow the Monod equation¹⁷, and are determined by the resource that is most limiting according to Liebig's 'law of the minimum'¹⁸:

$$\mu_i(R_1, \dots, R_k) = \min\left(\frac{r_i R_1}{K_{i1} + R_1}, \dots, \frac{r_i R_k}{K_{ik} + R_k}\right) \quad (3)$$

where r_i is the maximum specific growth rate of species i , K_{ji} is the half-saturation constant for resource j of species i , and $\min$ is the minimum function. This is a standard formulation used in numerous phytoplankton competition models^{6–10}.

When solved for equilibrium, this competition model predicts that the number of species cannot exceed the number of limiting resources. More precisely, there are k unknown resource availabilities in equation (1). Hence, in the generic case, the number of equilibrium solutions that satisfy equation (1) with $N_i > 0$ cannot

‡ Present address: Laboratory for Microbiology, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands.

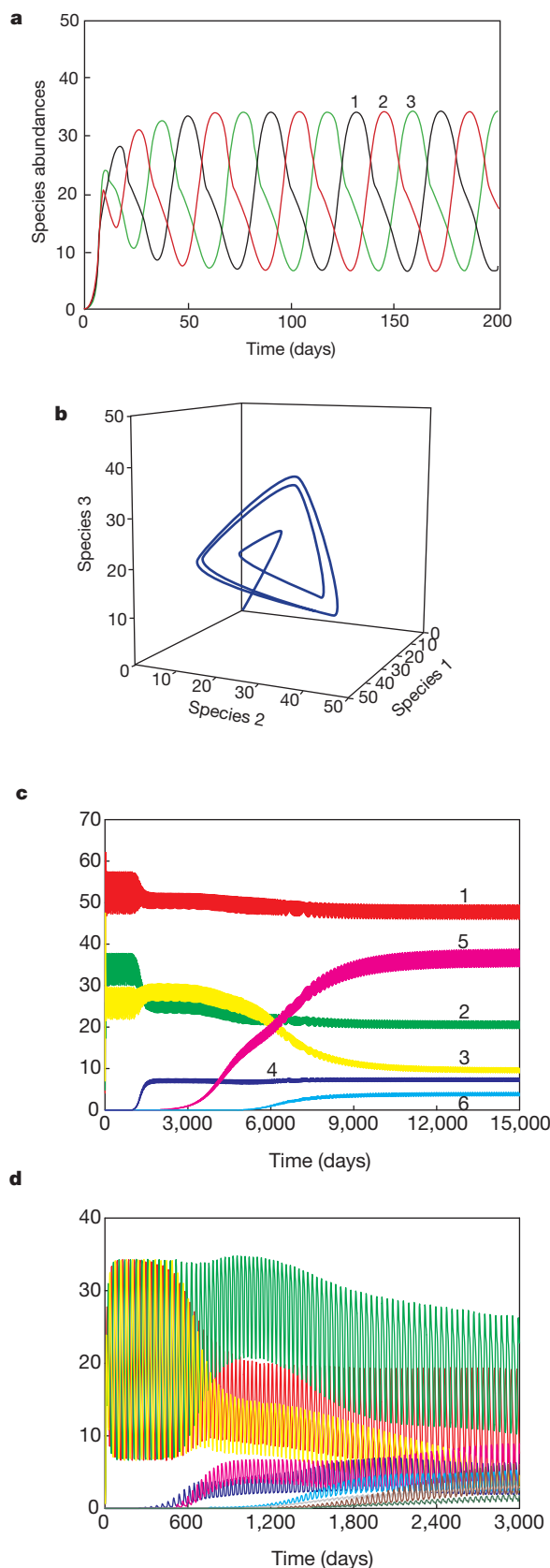


Figure 1 Oscillations on three resources. **a**, Time course of the abundances of three species competing for three resources. **b**, The corresponding limit cycle. **c**, Small-amplitude oscillations of six species on three resources. **d**, Large-amplitude oscillations of nine species on three resources.

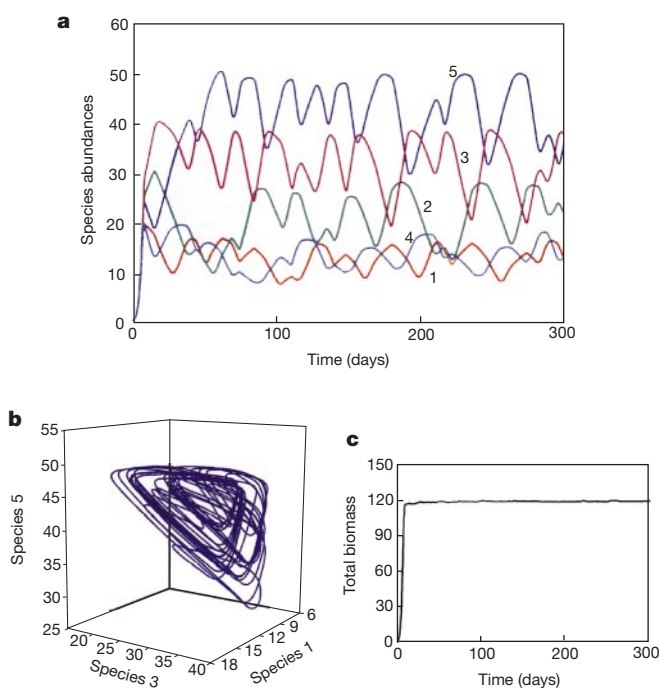


Figure 2 Chaos on five resources. **a**, Time course of the abundances of five species competing for five resources. **b**, The corresponding chaotic attractor. The trajectory is plotted for three of the five species, for the period from $t = 1,000$ to $t = 2,000$ days. **c**, Time course of total community biomass.

exceed k . The surplus of species should be competitively excluded. This leads to the so-called “principle of competitive exclusion”^{3–6}: at most $n \leq k$ species can coexist on k limiting resources. We note that this principle is based on equilibrium arguments. It assumes that competition leads to a stable species composition.

Several ways to circumvent the competitive exclusion principle and to explain the species diversity of planktonic communities have been proposed^{12,19–21}. These solutions usually invoke factors external to the phytoplankton, like selective predators, spatial heterogeneity, or temporal variability caused by fluctuating weather conditions. Here we develop a solution for the plankton paradox that does not invoke external factors. We consider a constant and homogeneous environment, and derive an explanation for biodiversity based on the dynamics of competition itself.

The dynamics of our competition model are well known for one or two limiting resources^{6–10}. In a constant environment, the system approaches a stable equilibrium. If all species are limited by the same resource, the strongest competitor displaces all other species and then approaches a monoculture equilibrium. If one species is limited by one resource, and another species by the other resource, then two species may stably coexist. Competition experiments with phytoplankton species support these predictions^{8,11–16}.

Natural phytoplankton communities, however, are frequently limited by more than two resources^{22–24}. What happens if the competition model is extended to three species and three resources? For certain species combinations, three-species competition generates sustained oscillations (Fig. 1a, b). This occurs if the species displace each other in a cyclic fashion. That is, species 1 is the better competitor for resource 1 but becomes limited by resource 2, species

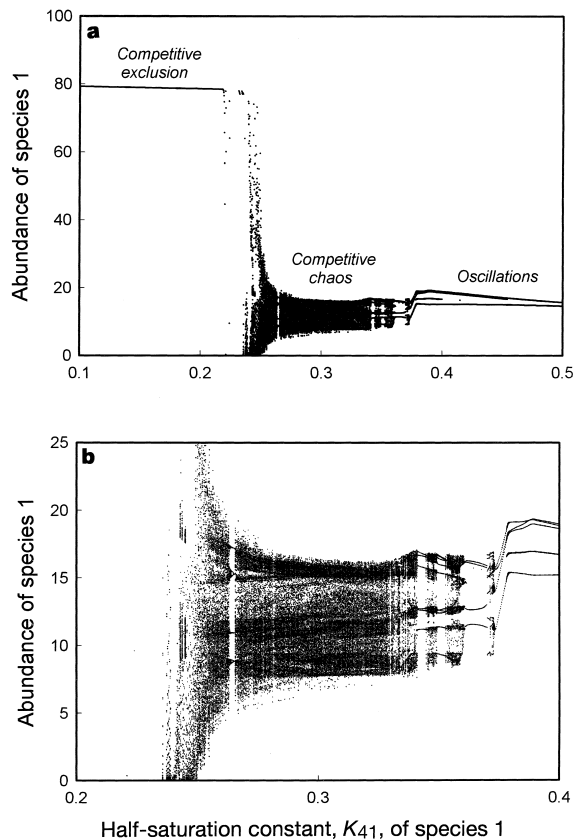


Figure 3 Bifurcation diagram, for five species competing for five resources. The graphs show the local minima and maxima of species 1, plotted during the period from $t = 2,000$ to $t = 4,000$ days, as a function of the half-saturation constant K_{41} . Part of **a** is magnified in **b**.

2 is the better competitor for resource 2 but becomes limited by resource 3, species 3 is the better competitor for resource 3 but becomes limited by resource 1, and so on. The amplitude of the species oscillations may range from small cycles (Fig. 1c) to large oscillations (Fig. 1d), depending on the precise parameter settings. We note that the oscillations are not generated by fluctuating weather conditions or other sources of external variability. The species oscillations are generated by the competition process itself.

Non-equilibrium conditions allow the coexistence of more species than limiting resources^{5,12,20}. Hence, it is conceivable that the oscillations generated by competition create an opportunity to increase species diversity. To test this idea, at $t = 1,000$ we added a fourth species to the model simulations (Fig. 1c). This fourth species is able to coexist on the oscillations generated by the three species already present. Also, a fifth species and a sixth species can be sustained. The amplitudes of the oscillations in Fig. 1c are so small that the oscillations would probably go unnoticed behind the noise of any real-world data set. Yet even these small-amplitude oscillations are apparently sufficient for the coexistence of six species on three resources. Similar results were obtained with large-amplitude oscillations (Fig. 1d): in the end, a total of nine species coexist on three resources.

Simulations revealed similar patterns with four limiting resources. For certain species combinations, competition for four resources generates oscillations. These oscillations allow the coexistence of many species on four resources (J.H. and F.J.W., unpublished results).

With five resources, many simulations show irregular species fluctuations (Fig. 2a). The pattern of species replacement never

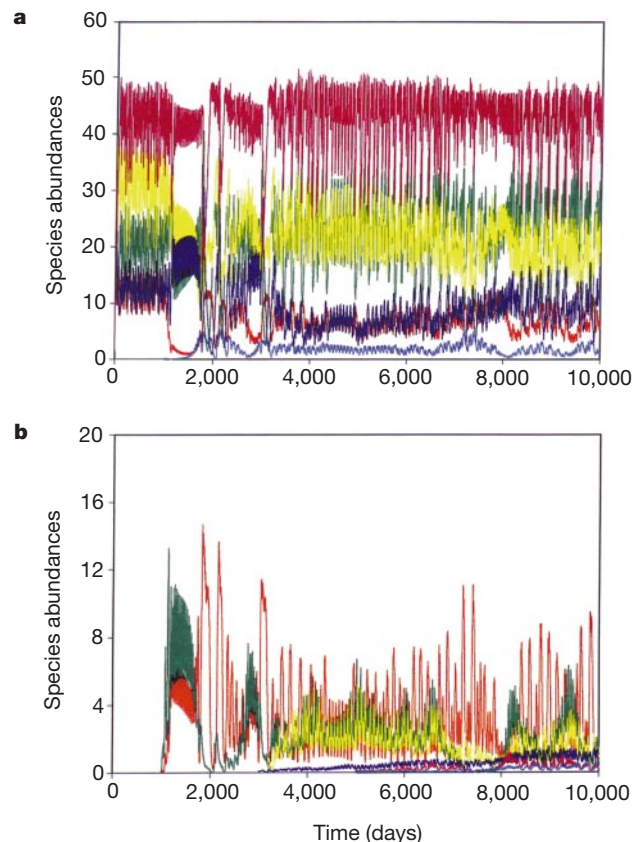


Figure 4 Competitive chaos and the coexistence of 12 species on five resources. **a**, The abundances of species 1–6; **b**, the abundances of species 7–12.

repeats itself. Each time one species tries to become dominant, there are several other species that invade. The species invade at different rates, and, hence, the abundances of the species continuously diverge. Yet all species abundances remain bounded because resources are limited. The continuous divergence of trajectories within a bounded region of phase space is a characteristic feature of chaos (Fig. 2b). In fact, the species dynamics show sensitive dependence on initial conditions. Extensive simulations reveal that trajectories that start with almost identical species abundances slowly diverge, and gradually become completely uncorrelated. The chaotic ups and downs of the individual species abundances go together with a near constancy of total community biomass (Fig. 2c). This supports the hypothesis^{25–27} that competition in high-diversity ecosystems may increase the variability at the species level while at the same time it may stabilize global ecosystem properties like total community biomass.

The bifurcation diagram in Fig. 3 illustrates how the model predictions depend on the parameter regime. We choose K_{41} , the half-saturation constant for resource 4 of species 1, as bifurcation parameter. Resource 4 is the resource that most limits the growth rate of species 1. If species 1 is a strong competitor for resource 4 ($K_{41} < 0.2$), species 1 excludes all other species (Fig. 3). If species 1 is a weak competitor for resource 4 ($K_{41} > 0.4$), competition leads to mild oscillations or stable coexistence. Competitive chaos occurs in the intermediate range, where species 1 is an “intermediate competitor” ($0.24 < K_{41} < 0.35$). Given the relatively broad parameter range that leads to chaos, it seems plausible that such competitors indeed occur in real-world plankton communities. We ran numerous additional simulations, with many parameter combinations

(not shown). This confirmed that the pattern in Fig. 3 is quite general: roughly speaking, competitive chaos occurs whenever each species is an intermediate competitor for the resources that most limit its growth rate. Chaotic fluctuations in species abundances allow the coexistence of many more species than limiting resources (Fig. 4).

The possibility that competition models may generate oscillations and chaos was already recognized in the mid 1970s^{28–30}. Also, it is well established that non-equilibrium conditions may favour species coexistence^{5,12,20}. What is new here is that we found both phenomena in a single competition model. Moreover, our findings do not stem from an artificially constructed model, but are based on one of the standard models of phytoplankton competition^{6–16}. We conclude that the biodiversity of plankton communities need not be explained by external factors, but can be based on the competition process itself. Once a plankton community is sufficiently complex to generate its own non-equilibrium dynamics, the number of coexisting phytoplankton species may greatly exceed the number of limiting resources, even in a constant and well-mixed environment. In this sense, the paradox of the plankton is essentially solved.

These findings have some wider implications that go beyond the plankton system studied here. First, within the biological realm, our explanation for planktonic biodiversity may serve as a conceptual model for the biodiversity of many other ecosystems as well. Second, our results demonstrate that competition, in its broadest sense, is not a simple process. Competitive systems may display highly dynamical phenomena, with continuous shifts and changes in species composition. Third, our results show that competition is not necessarily a destructive force. Competitive interactions that generate oscillations and chaos may allow the persistence of a great diversity of competitors on only a few limiting resources. □

Methods

Here we give the parameter values used in the model simulations. Simulations were based on equations (1) and (2), with specific growth rates according to equation (3). The model is parametrized for phytoplankton species, with a timescale expressed in days. We used $r_i = 1 \text{ d}^{-1}$ and $m_i = D = 0.25 \text{ d}^{-1}$ for all species. These are typical values for phytoplankton grown in chemostats^{8,11–16}. Initial conditions were $R_j = S_j$ and $N_i = 0.1 + i/100$ for all species i present at $t = 0$. Species added at a later time T started with $N_i = 0.1$ at $t = T$.

Half-saturation constants and resource contents for each resource j and species i are given below, using the compact notation of matrix algebra. Half-saturation constants, K_{ij} , are in matrix K . Resource contents, c_{ij} , are in matrix C . Different columns represent different species, and different rows represent different resources. Whereas our figures show coexistence for the first few thousand days, species continued to coexist in the simulations for more than 250,000 days.

Figures 1a, b use species 1–3 of Fig. 1d.

In Fig. 1c, $S_1 = 6$, $S_2 = 10$, $S_3 = 14$, and K and C are given by:

$$K = \begin{pmatrix} 1.00 & 0.90 & 0.30 & 1.04 & 0.34 & 0.77 \\ 0.30 & 1.00 & 0.90 & 0.71 & 1.02 & 0.76 \\ 0.90 & 0.30 & 1.00 & 0.46 & 0.34 & 1.07 \end{pmatrix}$$

$$C = \begin{pmatrix} 0.04 & 0.07 & 0.04 & 0.10 & 0.03 & 0.02 \\ 0.08 & 0.08 & 0.10 & 0.10 & 0.05 & 0.17 \\ 0.14 & 0.10 & 0.10 & 0.16 & 0.06 & 0.14 \end{pmatrix}$$

Species 1–3 start at $t = 0$, species 4 at $t = 1,000$, species 5 at $t = 2,000$, species 6 at $t = 5,000$.

In Fig. 1d, $S_1 = 10$, $S_2 = 10$, $S_3 = 10$ and K and C are given by:

$$K = \begin{pmatrix} 1 & 0.75 & 0.25 & 0.7 & 0.2 & 0.65 & 0.68 & 0.38 & 0.46 \\ 0.25 & 1 & 0.75 & 0.2 & 1.01 & 0.55 & 0.83 & 1.10 & 0.85 \\ 0.75 & 0.25 & 1 & 1.10 & 0.7 & 0.95 & 0.6 & 0.5 & 0.77 \end{pmatrix}$$

$$C = \begin{pmatrix} 0.10 & 0.20 & 0.15 & 0.05 & 0.01 & 0.40 & 0.30 & 0.20 & 0.25 \\ 0.15 & 0.10 & 0.20 & 0.15 & 0.30 & 0.35 & 0.25 & 0.02 & 0.35 \\ 0.20 & 0.15 & 0.10 & 0.25 & 0.05 & 0.20 & 0.40 & 0.15 & 0.10 \end{pmatrix}$$

Species 1–3 start at $t = 0$, species 4 at $t = 250$, species 5 at $t = 500$, species 6 at $t = 750$, species 7 at $t = 1,000$, species 8 at $t = 1,250$, species 9 at $t = 1,500$.

Figures 2 and 3 use species 1–5 of Fig. 4.

In Fig. 4, $S_1 = 6$, $S_2 = 10$, $S_3 = 14$, $S_4 = 4$, $S_5 = 9$, and K and C are given by:

$$K = \begin{pmatrix} 0.39 & 0.34 & 0.30 & 0.24 & 0.23 & 0.41 & 0.20 & 0.45 & 0.14 & 0.15 & 0.38 & 0.28 \\ 0.22 & 0.39 & 0.34 & 0.30 & 0.27 & 0.16 & 0.15 & 0.05 & 0.38 & 0.29 & 0.37 & 0.31 \\ 0.27 & 0.22 & 0.39 & 0.34 & 0.30 & 0.07 & 0.11 & 0.05 & 0.38 & 0.41 & 0.24 & 0.25 \\ 0.30 & 0.24 & 0.22 & 0.39 & 0.34 & 0.28 & 0.12 & 0.13 & 0.27 & 0.33 & 0.04 & 0.41 \\ 0.34 & 0.30 & 0.22 & 0.20 & 0.39 & 0.40 & 0.50 & 0.26 & 0.12 & 0.29 & 0.09 & 0.16 \end{pmatrix}$$

$$C = \begin{pmatrix} 0.04 & 0.04 & 0.07 & 0.04 & 0.04 & 0.22 & 0.10 & 0.08 & 0.02 & 0.17 & 0.25 & 0.03 \\ 0.08 & 0.08 & 0.08 & 0.10 & 0.08 & 0.14 & 0.22 & 0.04 & 0.18 & 0.06 & 0.20 & 0.04 \\ 0.10 & 0.10 & 0.10 & 0.10 & 0.14 & 0.22 & 0.24 & 0.12 & 0.03 & 0.24 & 0.17 & 0.01 \\ 0.05 & 0.03 & 0.03 & 0.03 & 0.03 & 0.09 & 0.07 & 0.06 & 0.03 & 0.03 & 0.11 & 0.05 \\ 0.07 & 0.09 & 0.07 & 0.07 & 0.07 & 0.05 & 0.24 & 0.05 & 0.08 & 0.10 & 0.02 & 0.04 \end{pmatrix}$$

Species 1–5 start at $t = 0$, species 6–8 at $t = 1,000$, species 9 and 10 at $t = 3,000$, species 11 and 12 at $t = 5,000$.

Received 17 June; accepted 20 September 1999.

- Wilson, E. O. *The Diversity of Life* (Belknap, Cambridge, Massachusetts, 1992).
- Hutchinson, G. E. The paradox of the plankton. *Am. Nat.* **95**, 137–145 (1961).
- Hardin, G. The competitive exclusion principle. *Science* **131**, 1292–1298 (1960).
- Phillips, O. M. The equilibrium and stability of simple marine biological systems. I. Primary nutrient consumers. *Am. Nat.* **107**, 73–93 (1973).
- Armstrong, R. A. & McGehee, R. Competitive exclusion. *Am. Nat.* **115**, 151–170 (1980).
- Grover, J. P. *Resource Competition* (Chapman and Hall, London, 1997).
- Leon, J. A. & Tumpson, D. B. Competition between two species for two complementary or substitutable resources. *J. Theor. Biol.* **50**, 185–201 (1975).
- Tilman, D. Resource competition between planktonic algae: an experimental and theoretical approach. *Ecology* **58**, 338–348 (1977).
- Hsu, S. B., Cheng, K. S. & Hubbell, S. P. Exploitative competition of micro-organisms for two complementary nutrients in continuous cultures. *SIAM J. Appl. Math.* **41**, 422–444 (1981).
- Huisman, J. & Weissing, F. J. Light-limited growth and competition for light in well-mixed aquatic environments: an elementary model. *Ecology* **75**, 507–520 (1994).
- Holm, N. P. & Armstrong, D. E. Role of nutrient limitation and competition in controlling the populations of *Asterionella formosa* and *Microcystis aeruginosa* in semicontinuous culture. *Limnol. Oceanogr.* **26**, 622–634 (1981).
- Sommer, U. Comparison between steady state and non-steady state competition: experiments with natural phytoplankton. *Limnol. Oceanogr.* **30**, 335–346 (1985).
- Sommer, U. Nitrate- and silicate-competition among Antarctic phytoplankton. *Mar. Biol.* **91**, 345–351 (1986).
- Van Donk, E. & Kilham, S. S. Temperature effects on silicon- and phosphorus-limited growth and competitive interactions among three diatoms. *J. Phycol.* **26**, 40–50 (1990).
- Rothhaupt, K. O. Laboratory experiments with a mixotrophic chrysophyte and obligately phagotrophic and phototrophic competitors. *Ecology* **77**, 716–724 (1996).
- Huisman, J., Jonker, R. R., Zonneveld, C. & Weissing, F. J. Competition for light between phytoplankton species: experimental tests of mechanistic theory. *Ecology* **80**, 211–222 (1999).
- Monod, J. La technique de culture continue, théorie et applications. *Ann. Inst. Pasteur (Paris)* **79**, 390–410 (1950).
- Von Liebig, J. *Die organische Chemie in ihrer Anwendung auf Agrikultur und Physiologie* (Friedrich Vieweg, Braunschweig, 1840).
- Richerson, P. J., Armstrong, R. & Goldman, C. R. Contemporaneous disequilibrium: a new hypothesis to explain the paradox of the plankton. *Proc. Natl Acad. Sci. USA* **67**, 1710–1714 (1970).
- Levins, R. Coexistence in a variable environment. *Am. Nat.* **114**, 765–783 (1979).
- Padisák, J., Reynolds, C. S. & Sommer, U. (eds) The intermediate disturbance hypothesis in phytoplankton ecology. *Hydrobiologia* **249**, 1–199 (1993).
- Sommer, U. Phytoplankton competition in Plußsee: a field test of the resource-ratio hypothesis. *Limnol. Oceanogr.* **38**, 838–845 (1993).
- Sterner, R. W. Seasonal and spatial patterns in macro- and micronutrient limitation in Joe Pool Lake, Texas. *Limnol. Oceanogr.* **39**, 535–550 (1994).
- Escaravage, V., Prins, T. C., Smaal, A. C. & Peeters, J. C. H. The response of phytoplankton communities to phosphorus input reduction in mesocosm experiments. *J. Exp. Mar. Biol. Ecol.* **198**, 55–79 (1996).
- May, R. M. *Stability and Complexity in Model Ecosystems* (Princeton Univ. Press, Princeton, 1974).
- Tilman, D. Biodiversity: population versus ecosystem stability. *Ecology* **77**, 350–363 (1996).
- Naeem, S. & Li, S. Biodiversity enhances ecosystem reliability. *Nature* **390**, 507–509 (1997).
- Gilpin, M. E. Limit cycles in competition communities. *Am. Nat.* **109**, 51–60 (1975).
- May, R. M. & Leonard, W. J. Nonlinear aspects of competition between three species. *SIAM J. Appl. Math.* **29**, 243–253 (1975).
- Smale, S. On the differential equations of species in competition. *J. Math. Biol.* **3**, 5–7 (1976).

Acknowledgements

J.H. thanks J. Roughgarden at Stanford University and L. J. Stal at the Center for Estuarine and Marine Ecology for their hospitality and for providing facilities to do this research. We thank U. Sommer for comments. J.H. was supported by the Netherlands Organization for Scientific Research (NWO), and the Earth and Life Sciences Foundation (ALW), which is subsidized by NWO.

Correspondence and requests for materials should be addressed to J.H. (e-mail: jef.huisman@chem.uva.nl).

The dynamics of chromosome evolution in birds and mammals

David W. Burt*, Charlotte Bruley*, Ian C. Dunn*, Cheryl T. Jones*, Anne Ramage*, Andy S. Law*, David R. Morrice*, Ian R. Paton*, Jacqueline Smith*, Dawn Windsor*, Alexei Sazanov†, Ruedi Fries† & David Waddington*

* Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK

† Lehrstuhl für Tierzucht, TUM, Freising, Weihenstephan, Munich D-85350, Germany

Comparative mapping, which compares the location of homologous genes in different species, is a powerful tool for studying genome evolution¹. Comparative maps suggest that rates of chromosomal change in mammals can vary from one to ten rearrangements per million years^{1–4}. On the basis of these rates we would expect 84 to 600 conserved segments in a chicken comparison with human or mouse. Here we build comparative maps between these species and estimate that numbers of conserved segments are in the lower part of this range. We conclude that the organization of the human genome is closer to that of the chicken than the mouse and by adding comparative mapping results from a range of vertebrates, we identify three possible phases of chromosome evolution. The relative stability of genomes such as those of the chicken and human will enable the reconstruction of maps of ancestral vertebrates.

To examine the dynamics of chromosome rearrangement during vertebrate evolution, we constructed comparative maps of chicken with human and mouse (see Supplementary Information), species that share a common ancestor 300 million years ago. We have used the recommendations of the Human Genome Organisation (HUGO) Comparative Genome Organisation for definitions of gene homologies, orthologies and conserved segments¹. Over 2,000 orthologous genes ('orthologues') have already been mapped in a mouse–human comparison² and these define 195 autosomal conserved segments. Given the large number of genes, we

have assumed that the actual number is close to this total. In the chicken, we mapped 223 genes which define 81 autosomal conserved segments in the human comparison and 100 in the mouse comparison. We predict the total number of autosomal conserved segments in the chicken–human comparison to be 96 (95% confidence limits: lower 79; upper 117) and in the chicken–mouse (ref. 5) comparison to be 152 (118, 198). These totals are the sum of the number of autosomes in the last common ancestor (assumed to be 24; ref. 6) and the number of chromosome rearrangements since divergence³. Therefore we predict 72 (55, 93) chromosome rearrangements for the chicken–human comparison; this is less than for mouse–chicken (128, (94, 174)) and mouse–human (at least 171). We conclude that the organization of the human genome is closer to that of the chicken than the mouse.

A relative rate test⁷ was used to compare the rates of chromosomal change in human and mouse lineages, using the chicken lineage as an outgroup. The rate in the mouse lineage was twice as fast as that in the human. This difference is supported if we compare the sizes of conserved segments that are common to both our chicken–mouse and chicken–human comparative maps. Of 14 segments that differ in size, 13 are larger in the chicken–human comparison (sign-test $P < 0.001$). Rates for human and mouse lineages are 0.58 and 1.14 rearrangements per million years, respectively, since divergence 100 Myr ago. There is no outgroup for chicken, so it is not possible to estimate the number of rearrangements in this lineage; however, the maximum number compatible with the above upper limits for human and mouse is 48. This gives a maximum rate of 0.16 rearrangements per million years since divergence 300 Myr ago, which is considerably less than the minimum estimate for the mouse lineage (0.9 per Myr).

To reveal the global pattern of chromosome evolution we have combined results from genetic mapping and chromosome painting (Zoo-FISH) with vertebrate phylogeny^{6–10}. A summary is shown in Fig. 1, from which we identify three phases of chromosome evolution. In phase I (100–300 Myr ago), the rate of chromosome rearrangement was slow, less than 0.2 per million years in both avian and mammalian lineages. In phase II (65–100 Myr ago), the rate increased to over 1.1 per million years in both mouse and non-rodent lineages. Recent comparisons of rat and human¹¹ suggest

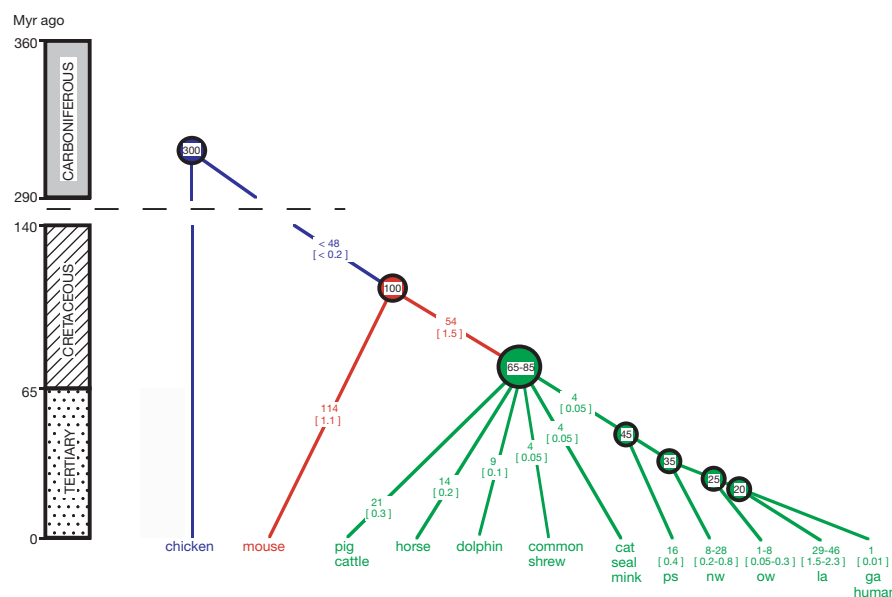


Figure 1 Dynamics of chromosome evolution in birds and mammals. Circles represent estimated times (in Myr) of divergence of common ancestors. The three phases of chromosomal evolution discussed in the text are shown in blue (phase I), red (phase II) and green (phase III). The estimated number of chromosome rearrangements is shown along

each lineage, with the rates of chromosomal rearrangement per million years (in brackets): ps, prosimians (lemur); nw, New World monkeys (six species); ow, Old World monkeys (four species); la, lesser apes (four species); ga, great apes (six species). For more details see Methods and Supplementary Information.

that rates of chromosome change may have been high in mice (1.1 per Myr), rats (0.7 per Myr) and possibly other rodents. In phase III (during the past 65–85 million years) the rate has been variable in non-rodent mammals, with the slowest rate (less than 0.1 per Myr) for human, carnivores and common shrew, a higher rate (0.1–0.3 per Myr) for pig, cattle, horse and dolphin, and the highest rate in the lesser apes (1.5–2.3 per Myr). These rates of chromosome evolution differ from earlier estimates¹², measured from changes in chromosome and chromosome arm number, averaged over genera within major taxonomic groups. The differences we find within specific lineages cannot be detected by this measure.

How can we explain these patterns of chromosome evolution, and how are they explained by current models of genome evolution? Many chromosome changes are likely to be deleterious and will therefore be rapidly lost by natural selection; they will thus not contribute to chromosome diversity. Some slightly deleterious mutations are associated with reduced fitness, for example, owing to reduced fertility caused by abnormal segregation of chromosomes during meiosis. These are unlikely to be fixed unless the mutation rate is high or the population size is small.

When the population size becomes small (such as after a natural disaster or at the time of speciation), then genetic drift will predominate and chromosomal rearrangements with reduced fitness will have a greater chance of fixation¹². As the population expands, natural selection becomes increasingly influential. Then the rate of chromosomal evolution will diminish to a basal rate of change, possibly representing rare, advantageous mutations. Under natural selection, these mutations will be fixed at a rate that is largely controlled by how much the environment changes, which would depend on chronological time. This rate will tend to be equal in all lineages. Fluctuations in population size may have been a major factor leading to the wide variation in the rates of chromosome rearrangement in non-rodent mammals in phase III (Fig. 1). Such changes would be expected 65–85 Myr ago, a time of rapid speciation in mammals.

When the mutation rate is the primary cause of differences in the rates of chromosomal rearrangement between lineages, then mutation mechanisms must differ. One possible mechanism is chromosome mispairing (and subsequent translocation) caused by homologous sequences at different sites in the genome⁶. Then changes in the frequency of mispairing may alter the mutation rate. The frequency of at least three types of homologous sequence differs between mammalian and avian genomes. Dispersed repetitive sequences make up more than 50% of a mammalian genome but less than 15% of a bird genome. Also, mammals have larger gene families than birds, often with many pseudogenes. Multiple copies of retroviral elements are more frequent in mammals, and these may play a role in some rearrangements¹³. Consequently, the potential for chromosome mispairing may be low in birds and this may explain the low rate of chromosomal rearrangement within the chicken lineage. It may also have led to a low rate in ancestral mammals, 100–300 Myr ago (phase I, Fig. 1). Perhaps the accumulation of homologous sites in rodent and non-rodent lineages increased the rate of chromosome change in phase II (Fig. 1). Our findings on rates of chromosomal rearrangement show similarities to the comparison of nucleotide substitution rates^{14,15} between mouse and human. Both rates appear to be more dependent on generation time than chronological time, with the mouse rates being at least twice those found in human during the last 100 million years. This is an inevitable consequence of different generation times between mouse and human, if we assume equal mutation rates per DNA replication¹⁶. Differences in generation time will also contribute to the variation in chromosomal evolution within non-rodent mammals in phase III (Fig. 1).

Conservation of genome organization throughout the vertebrate kingdom has important practical and evolutionary implications. First, we can be optimistic about the use of vertebrate comparative

maps to predict candidate genes for phenotypes mapped in species as diverse as chicken and human. Second, we can now start to use comparative gene maps derived from selected species for the systematic reconstruction of ancestral vertebrate genomes. These will enable us to establish the dynamics of chromosomal reorganization. This will be made possible by the relative stability of some of the genomes being mapped, such as chicken and human, spanning 300 million years of vertebrate evolution. □

Methods

Chicken maps are recorded in Arkdb-CHICK (<http://www.ri.bbsrc.ac.uk/chickmap/>), and at World-Wide Web sites in Wageningen (<http://www.zod.wau.nl/vf/chickensite/chicken.html>) and East Lansing (<http://poultry.mph.msu.edu/chickmap.html>). Human and mouse maps were taken from GDB (<http://gdbwww.gdb.org/gdb>), Gene Map'98 (<http://www.ncbi.nlm.nih.gov>) and MGD (<http://www.informatics.jax.org>, March 1999).

The East Lansing map¹⁷ was used as a reference on which to map all genes. For each chromosome, common markers were linked by a non-parametric regression of genetic distances from other crosses onto East Lansing distances, using cubic splines¹⁸. The degree of smoothing was selected by minimizing the Akaike information criterion¹⁹. When there were two or three markers, a linear or quadratic regression was fitted, as appropriate. To estimate the length of the chicken genome, map lengths were converted to genetic distances using the Kosambi mapping function²⁰ and adjusted for failure to sample telomeric regions²¹. From this approach the total genetic length was estimated to be 3900 cM.

The method of Waddington *et al.*⁵ was used to estimate the underlying total number of conserved segments, using as data both the counts of genes in a syntenic block and distances between them. Only genes mapped at random (134 out of 223 mapped genes), that is, with no *a priori* knowledge of map position⁵, were included in our analysis. Syntenic blocks defined by a single gene were accepted only when sequence comparisons could be made to minimize orthology errors. Many were confirmed by linkage to other genes not mapped at random or to a confirmed orthology in a human–mouse comparison. See Supplementary Information for a detailed summary of the observed/predicted numbers of conserved segments in the chicken comparisons.

Divergence times for birds and mammals were obtained from published sources^{8–10}, but definite dates were difficult to obtain owing to conflicts between sequence and fossil data¹⁰. In non-rodent mammals this is acknowledged to Fig. 1 by a star phylogeny radiating from a common ancestor 65–85 Myr ago^{8–10}. Using values at the extremes of this range alters our conclusions about rates of change very little. Phylogenetic studies generally, but not always, place mice before the divergence of all other mammals⁹. We estimate 58 chromosome rearrangements in the human lineage, since divergence from a common ancestor with mice, but only 8 since divergence from the cat⁴. The difference is consistent with mice being placed in the more ancient branch, 100 Myr ago¹⁰.

The resolution of Zoo-FISH is limited to 10 Mb (ref. 22) and so the observed number of conserved segments may be underestimated. However, assuming random chromosome breaks throughout the genome⁵, we would expect few small segments in lineages with few chromosome rearrangements and therefore this bias should be small. Most Zoo-FISH data used in the construction of Fig. 1 were based on comparisons with human and so it was not possible to calculate the exact number of chromosome rearrangements in each lineage by solving simultaneous equations. However, since the number of rearrangements between human and cat was small⁴, we assumed that the rate was equal in these two lineages. With this approximation, the numbers of rearrangements in the other lineages were derived by subtraction.

Received 5 May; accepted 30 September 1999.

- Andersson, L. *et al.* The first international workshop on comparative genome organisation. *Mamm. Genome* **7**, 717–734 (1997).
- Davison, M. T., Bradt, D. W., Merriam, J. J., Rockwood, S. F. & Eppig, J. T. The mouse gene map. *Inst. Lab. Anim. Res. J.* **39**, 96–131 (1998).
- Nadeau, J. H. & Taylor, B. A. Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc. Natl. Acad. Sci. USA* **81**, 814–818 (1984).
- O'Brien, S. J., Wienberg, J. & Lyons, L. A. Comparative genomics: lessons from cats. *Trends Genet.* **13**, 393–399 (1997).
- Waddington, D., Springbett, A. J. & Burt, D. W. A chromosome based model to estimate the number of conserved segments between pairs of species from comparative genetic maps. *Genetics* (in the press).
- Morizot, D. C. Reconstructing the gene map of the vertebrate ancestor. *Anim. Biotech.* **5**, 113–122 (1994).
- Sarich, V. M. & Wilson, A. C. Immunological time scale for hominid evolution. *Science* **158**, 1200–1203 (1994).
- Novacek, M. J. Mammalian phylogeny: shaking the tree. *Nature* **356**, 121–125 (1992).
- de Jong, W. W. Molecules remodel the mammalian tree. *Trends Ecol. Evol.* **13**, 270–275 (1998).
- Kumar, S. & Hedges, B. A molecular time-scale for vertebrate evolution. *Nature* **392**, 917–920 (1998).
- Watanabe, T. K. *et al.* A radiation hybrid map of the rat genome containing 5,255 markers. *Nature Genet.* **22**, 27–36 (1999).
- Bush, G. L., Case, S. M., Wilson, A. C. & Patton, J. L. Rapid speciation and chromosomal evolution in mammals. *Proc. Natl. Acad. Sci. USA* **74**, 3942–3946 (1977).
- O'Neill, R. J. W., O'Neill, M. J. & Graves, J. A. M. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393**, 68–72 (1998).
- Li, W.-H., Ellsworth, D. L., Krushkal, J., Chang, B. H.-J. & Hewett-Emmett, D. Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylogenet. Evol.* **5**, 182–187 (1996).

15. Ohta, T. An examination of the generation-time effect on molecular evolution. *Proc. Natl. Acad. Sci. USA* **90**, 10676–10680 (1993).
16. Vogel, F., Kopun, M. & Rathenberg, R. in *Molecular Anthropology* (eds Goodman, M. & Tashian, R. E.) 13–33 (Plenum, New York, 1976).
17. Burt, D. W. & Cheng, H. H. The chicken gene map. *Inst. Lab. Anim. Res. J.* **39**, 229–236 (1998).
18. Hastie, T. J. & Tibshirani, R. J. *Generalised Additive Models* (Chapman and Hall, London, 1990).
19. Hurvich, C. M., Simonoff, J. S. & Tsai, C.-L. Smoothing parameter selection in non-parametric regression using an improved Akaike information criterion. *J. R. Statist. Soc.* **60**, 271–293 (1998).
20. Kosambi, D. D. The estimation of map distance from recombination values. *Ann. Eugen.* **12**, 172–175 (1944).
21. Morton, N. E. Parameters of the human genome. *Proc. Natl. Acad. Sci. USA* **88**, 7474–7476 (1991).
22. Wienberg, J. & Stanyon, R. Comparative chromosome painting of primate genomes. *Inst. Lab. Anim. Res. J.* **39**, 77–91 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London Editorial office of Nature.

Acknowledgements

We thank members of the EC ChickMAP project and H. Cheng for providing data prior to publication, and J. Burt for helpful comments. Genome research at the Roslin Institute is supported by the Ministry of Agriculture, Fisheries and Food, the Biotechnology and Biological Sciences Research Council and the Commission of the European Communities.

Correspondence and requests for materials should be addressed to D.W.B. (e-mail: Dave-Burt@bbsrc.ac.uk).

Large-scale analysis of the yeast genome by transposon tagging and gene disruption

Petra Ross-Macdonald*, Paulo S. R. Coelho*, Terry Roemer*, Seema Agarwal*, Anuj Kumar*, Ronald Jansen†, Kei-Hoi Cheung‡, Amy Sheehan*, Dawn Symoniatis*, Lara Umansky*, Matthew Heidman*, F. Kenneth Nelson*, Hiroshi Iwasaki*, Karl Hager\$, Mark Gerstein†, Perry Miller‡, G. Shirleen Roeder* & Michael Snyder*†

* Department of Molecular, Cellular and Developmental Biology, Yale University, PO Box 208103, New Haven, Connecticut 06520-8103, USA

† Department of Molecular Biophysics and Biochemistry, Yale University, PO Box 208114, New Haven, Connecticut 06520-8114, USA

‡ Center for Medical Informatics, Department of Anesthesiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

\$ Keck Foundation Biotechnology Resource Laboratory, 295 Congress Avenue, Yale University, New Haven, Connecticut 06520, USA

Economical methods by which gene function may be analysed on a genomic scale are relatively scarce. To fill this need, we have developed a transposon-tagging strategy for the genome-wide analysis of disruption phenotypes, gene expression and protein localization, and have applied this method to the large-scale analysis of gene function in the budding yeast *Saccharomyces cerevisiae*. Here we present the largest collection of defined yeast mutants ever generated within a single genetic background—a collection of over 11,000 strains, each carrying a transposon inserted within a region of the genome expressed during vegetative growth and/or sporulation. These insertions affect nearly 2,000 annotated genes, representing about one-third of the 6,200 predicted genes in the yeast genome^{1,2}. We have used this collection to determine disruption phenotypes for nearly 8,000 strains using 20 different growth conditions; the resulting data sets were clustered to identify groups of functionally related genes. We have also identified over 300 previously non-annotated open reading frames and analysed by indirect immunofluorescence over 1,300 transposon-tagged proteins. In total, our study encompasses over 260,000 data points, constituting the largest functional analysis of the yeast genome ever undertaken.

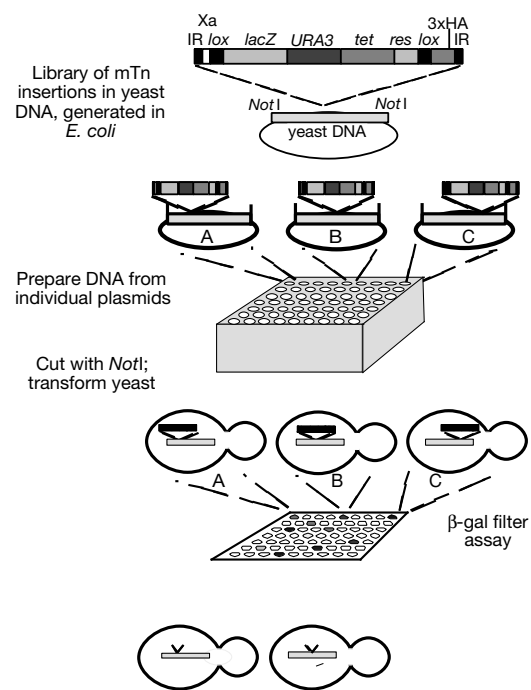


Figure 1 The mTn insertion project. Most steps were performed using a Robbins Hydra 96-channel dispenser; all strains are maintained in a 96-well format.

The ability to sequence entire genomes has resulted in an abundance of raw sequence data³; however, relatively few methods exist to assess gene function on a genomic scale^{4–8}. We have developed a transposon-based method for the large-scale accumulation of expression, phenotypic and protein localization data in yeast without bias towards previously annotated genes. Our approach utilizes a multipurpose minitransposon (mTn) derived from the bacterial transposable element Tn3 (ref. 9). The minitransposon mTn-3xHA/lacZ (Fig. 1) contains a *lacZ* reporter gene lacking an initiator methionine and upstream promoter sequence. Introduction of this transposon into yeast results in production of β-galactosidase (β-gal) if the mTn is present within a transcribed and translated region of the genome, typically corresponding to an in-frame fusion of *lacZ* to the yeast protein-coding sequence. Additionally, mTn-3xHA/lacZ contains a *lox* site near each Tn3 end; adjacent to one *lox* site is DNA encoding three copies of a haemagglutinin (3xHA) epitope tag. Production of the Cre recombinase in yeast containing this minitransposon results in recombination of the *lox* sites, reducing the mTn to a 274 base pair (bp) element (the HAT tag)⁹. As five bases of genomic DNA are duplicated during Tn3 insertion, the net result is a 279-bp insertion encoding a 93-codon open reading frame (ORF) which includes the 3xHA sequence. When the mTn's *lacZ* reporter has been fused in-frame to a yeast coding region, creation of the HAT tag allows production of a full-length, epitope-tagged protein from that gene.

We have used mTn-3xHA/lacZ in conjunction with high throughput methods to generate a large collection of yeast strains, each containing an insertion at a known location within the genome (Fig. 1). Briefly, a yeast genomic DNA library was mutagenized in *Escherichia coli* with mTn-3xHA/lacZ. Using a 96-well format, individual plasmids were prepared, digested with *NotI* and transformed into a diploid yeast strain¹⁰. By homologous recombination, each fragment should integrate at its corresponding genomic locus, thereby replacing its genomic copy. To verify this process, polymerase chain reaction (PCR) analysis was performed on 48 independent yeast transformants, each carrying one of six different mTn insertion alleles. Using a primer in the mTn and a primer external to

flanking genomic DNA, we confirmed that the allele had substituted properly in 47 out of 48 cases.

Transformants carrying the mTn-borne *URA3* marker were selected and assayed for β -gal activity after growth on rich medium and sporulation medium¹⁰. Strains were also treated to induce the Cre recombinase. Following this Cre-mediated excision event, we used *Ura*[−] colonies for immunolocalization of epitope-tagged proteins with anti-HA antibodies. The bacterial clone that gave rise to each yeast strain in the collection was also stored. Plasmid DNA from these bacteria allowed transformation of insertion alleles into a haploid strain for analysis of disruption phenotypes.

Using this procedure, we have carried out 92,544 plasmid preparations and yeast transformations, allowing us to identify 11,232 strains containing *lacZ* fusions expressed during vegetative growth; the precise site of mTn insertion within the yeast genome has been determined by DNA sequence analysis for 6,358 strains. As expected, most of the strains within our collection contain transposon insertions affecting annotated ORFs. Of these 6,358 insertion alleles, 5,442 lie in or within 200 bp of an annotated ORF. In total, these insertions affect 1,917 ORFs distributed over all 16 chromosomes of the yeast genome (Fig. 2).

Most of these transposon insertions represent in-frame fusions to coding sequence; we have identified in-frame fusion events in 1,346 annotated ORFs (Fig. 2). However, reproducibly high levels of β -gal activity were also found in around 480 strains carrying a transposon located within an annotated ORF, but in either the +1 or −1 reading frame. In a study of 287 such insertions, we found that an alternative in-frame start codon is lacking in 75% of the sequences, indicating that a frameshift may occur during translation of the annotated ORF. Translational frameshifting within the yeast genome may, therefore, occur more frequently than has been reported.

Figure 2 Distribution and phenotypic analysis of 1,917 mTn-mutagenized ORFs within the *S. cerevisiae* genome. Each chromosome is represented as a linear map with start and end coordinates indicated to the left and right, respectively. Mutagenized ORFs are shown as proportionally sized boxes within either the Watson (upwards) or Crick (downwards) strand of each chromosome. Common gene names have been used when available; otherwise, ORFs are designated by the third through sixth characters of their standardized names. All in-frame mTn3 fusions are indicated with asterisks. A subset of putative NORFs located within intergenic regions are shown in their corresponding chromosomal positions; these NORFs were identified by multiple mTn insertion events. NORFs included have been designated N400 through N442 (see ref. 4). Macroarray analysis has identified disruption phenotypes associated with 407 ORFs; these ORFs have been colour-coded as indicated (see Table 1). Complete data sets can be accessed at <http://ycmi.med.yale.edu/ygac/triples.htm>.

In analysing mTn insertion events, we have identified a large number of previously non-annotated ORFs (NORFs) within the yeast genome. When annotating the *S. cerevisiae* genome, an arbitrary lower size limit of 100 amino acids was adopted as convention unless additional information suggested otherwise. Also, ORFs of more than 100 codons that fully or partially overlapped even longer ORFs were usually not annotated^{4,11}. Such NORFs are, however, capable of being translated, as indicated in previous studies^{4,8}. Using an arbitrary lower limit of 50 codons, we find that 328 of 6,358 mTn insertions represent in-frame fusions to NORFs. These NORFs range in size up to 247 codons. A detailed analysis of 229 such insertions revealed three classes: 52% are in NORFs that fully or partially overlap an annotated ORF in the antisense direction; 15% fall in NORFs that overlap an annotated ORF in the same orientation but in a different frame; and 33% of these insertions occurred in NORFs previously classified as intergenic regions. Many NORFs were identified by multiple mTn insertion events. For example, a 124-codon NORF in the antisense

Table 1 Phenotypes scored by macroarray analysis

Assay/growth conditions	Number*	Colour code†	Abbreviation‡
YPD + 8 mM caffeine	27	Purple	Caff
Cycloheximide hypersensitivity: YPD + 0.08 μ g ml ^{−1} cycloheximide at 30 °C	28†	Pink	Cyc ^S
White/red colour on YPD	39	Yellow	W/R
YPGlycerol	54	Yellow	YPG
Calcofluor hypersensitivity: YPD + 12 μ g ml ^{−1} calcofluor at 30 °C	66	Purple	Calc ^S
YPD + 46 μ g ml ^{−1} hygromycin at 30 °C	136	Purple	Hyg
YPD + 0.003% SDS	109	Purple	SDS
Benomyl hypersensitivity: YPD + 10 μ g ml ^{−1} benomyl	67	Green	Ben ^S
YPD + 5-bromo-4-chloro-3-indolyl phosphate at 37 °C	35	Purple	BCIP
YPD + 0.001% methylene blue at 30 °C	12	Purple	MB
Benomyl resistance: YPD + 20 μ g ml ^{−1} benomyl	11†	Green	Ben ^R
YPD at 37 °C	29	Cyan	YPD ³⁷
YPD + 2 mM EGTA	30	Black	EGTA
YPD + 0.008% MMS	16	Pink	MMS
YPD + 75 mM hydroxyurea	21	Pink	HU
YPD at 11 °C	20	Cyan	YPD ¹¹
Calcofluor resistance: YPD + 0.3 μ g ml ^{−1} calcofluor at 30 °C	4†	Purple	Calc ^R
Cycloheximide resistance: YPD + 0.3 μ g ml ^{−1} cycloheximide	2†	Pink	Cyc ^R
Hyperhaploid invasive growth mutants	25	Orange	HHIG
YPD + 0.9 M NaCl	13	Black	NaCl

All tests were carried out at room temperature (~24 °C) unless otherwise noted.

* The total number of ORFs conferring a given phenotype upon disruption is indicated.

† Data sets are incomplete for assays of cycloheximide hypersensitivity (5,760 transformants screened), cycloheximide resistance (3,456 transformants screened), calcofluor resistance (2,208 transformants screened) and benomyl resistance (5,760 transformants screened).

‡ Colours correspond to those shown in Fig. 2.

§ The abbreviations for each growth condition are used as labels in Fig. 4.

The tree to the left of the table was generated by cluster analysis and indicates growth conditions that were found to result in most similar phenotypes for the entire data set (see also Fig. 4).

region of the rDNA has more than 50 insertions. In addition, we have observed mutant phenotypes resulting from transposon insertions within putative NORFs: for example, mTn insertion at base number 198,816 of chromosome XII (identifying a putative intergenic NORF of 86 amino acids) results in hypersensitivity to the microtubule-depolymerization drug benomyl. Furthermore, three NORFs encode proteins exhibiting a distinct pattern of subcellular localization. Transposon insertion at base number 1,236,754 of chromosome IV identifies a NORF of 97 amino acids; using the mTn-encoded epitope tag, we have localized the encoded protein to the nucleus. These findings substantiate the biological importance of NORFs, raising the possibility that they encode previously uncharacterized proteins. Complete data sets for all strains carrying mTn insertions identifying putative NORFs may be accessed from our web site at <http://ycmi.med.yale.edu/ycac/triples.htm>.

To analyse phenotypes on a genomic scale, we transformed 7,680 mTn-insertion alleles into a haploid strain. Most transformants were viable. In the 14.1% (1,082) that were inviable, the transposon typically affected a gene known to be required for cell viability. Gene functions associated with these essential genes may also be analysed using our multipurpose minitransposon. By *cre-lox*-mediated reduction of mTn-3xHA/*lacZ*, we have generated viable mutants containing in-frame insertions of HAT-tags within essential genes. We estimate that 30–45% of small insertion alleles in essential genes retain gene function.

Transformed haploid strains were scored for 20 phenotypes after

growth under the test conditions indicated in Table 1 and Fig. 2. Transformants were screened on a large scale using 'phenotypic macroarrays' compatible with 96-well formats. Yeast strains grown in 96-well arrays were transferred to test medium using a 96-pin tool; 576 strains were simultaneously screened for a given phenotype (Fig. 3).

Transposon insertions in 407 genes resulted in a phenotype significantly different from the wild type (excluding those insertion alleles causing inviability). Many mutant alleles caused phenotypes that might be expected. For example, strains carrying insertions in *TUB3* (ref. 12), *CIK1* (ref. 13), and *CIN1* (ref. 14)—genes known to be involved in microtubule functions—are sensitive to benomyl. In other cases, insertions within characterized genes resulted in unexpected phenotypes, indicating that these genes may have previously unrecognized roles or affect multiple processes indirectly. For example, the serine/threonine phosphatase Rts1p (ref. 15) is involved in the control of stress-related responses and cell-cycle regulation, whereas *YMA5* (ref. 16) encodes a hydrophilic subunit of a vacuolar H^+ -ATPase; mutations affecting both these ORFs caused sensitivity to the cell wall-binding dye calcofluor, indicative of genes involved in cell wall biogenesis¹⁷. Other insertions caused mutant phenotypes in a variety of apparently unrelated assays: strains containing insertions in genes for nuclear pore complex proteins and for certain transcription factors were sensitive to hydroxyurea, benomyl and calcofluor. We speculate that genes whose mutations result in highly pleiotropic phenotypes are probably involved in

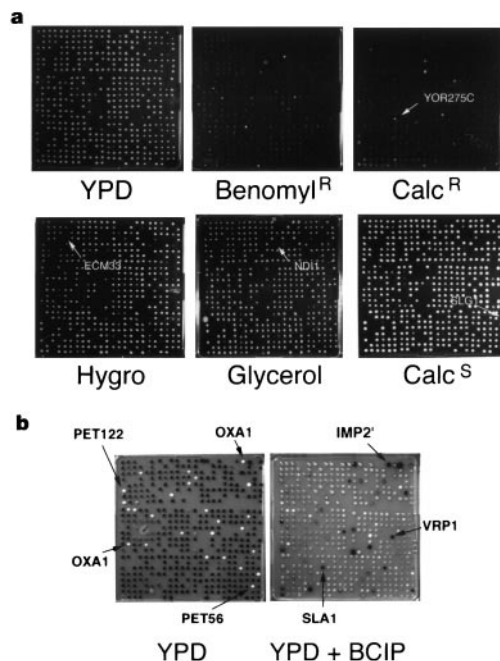


Figure 3 Phenotypic macroarray analysis. **a**, Examples of 21 cm × 21 cm macroarray plates scoring growth on test media: YPD, YPD supplemented with 20 $\mu\text{g ml}^{-1}$ benomyl, YPD with 66.7 $\mu\text{g ml}^{-1}$ calcofluor, YPD with 46 $\mu\text{g ml}^{-1}$ hygromycin, YPGlycerol, YPD supplemented with 12 $\mu\text{g ml}^{-1}$ calcofluor. Arrows indicate strains mutated for genes functioning in cellular respiration (*ND1*) or cell-wall biogenesis (*ECM33*, *SLG1*, *YOR275C*)¹⁷. **b**, Macroarray analysis of yeast mutants deficient in oxidative phosphorylation and cell-wall maintenance. Mitochondrial mutants unable to carry out oxidative phosphorylation were identified as white (rather than red) colonies on YPD medium; red pigment formation within *ade2* mutant strains requires oxidative phosphorylation²⁸. Representative respiratory genes identified within our screen are labelled with arrows (*PET122*, *PET56*, *OXA1*). Genes functioning in cell-wall maintenance were characterized through macroarray analysis of mutants grown on YPD overlaid with agar containing BCIP. BCIP is a chromogenic substrate of alkaline phosphatase, which when released from vacuoles during cell lysis cleaves BCIP, thereby staining lysed mutants blue. Arrows highlight a sample of genes (*IMP2*, *VPR1*, *SLA1*)¹⁷ involved in cell-wall maintenance.

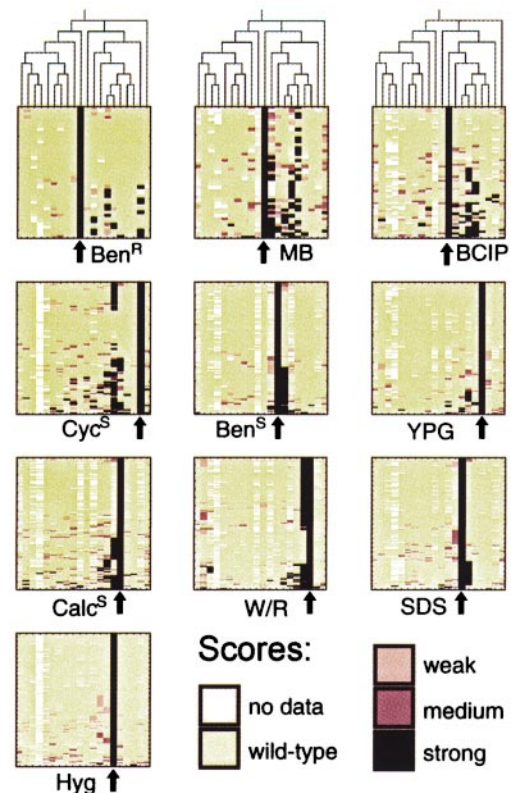


Figure 4 Graphical representation of clustered phenotypic data. Examples of data clusters exhibiting a predominant phenotype are shown (all 21 clusters may be viewed at <http://bioinfo.mbb.yale.edu/genome/phenotypes>). Vertical columns represent growth conditions; each horizontal row represents a different transformant. Transformants in each cluster are sorted from top to bottom according to the distance from the cluster average (centroid). The abbreviation below each square indicates the predominant phenotype of each cluster (see Table 1). The arrow below each square points to the column representing this predominant phenotype. In addition to clustering transformant phenotypes (the rows of the data set), we have also clustered vertical data columns (representing assayed growth conditions). The tree generated by column clustering is shown above the first row of clusters (as well as to the left of Table 1).

Table 2 Observed immunofluorescence patterns for in-frame HAT fusions

Pattern	Number*
Discrete	
Nuclear	
general	41
granular	64
nucleolus	22
Nuclear rim/endoplasmic reticulum	29
Mitochondrial	37
Spindle pole body/microtubules	5
Cell periphery	11
Cytoplasmic patch/dots	10
Cell neck	2
Vacuole	2
General cytoplasmic	
Uniform, finely speckled	3
Granular, fibrous	189
Background	925
Total	1,340

* All strains exhibiting discrete cellular or cytoplasmic localizations were tested at least twice. Complete data sets for all tested clones can be accessed at <http://ycmi.med.yale.edu/ycac/triples.htm>.

general cellular processes. Finally, many strains carrying insertions within uncharacterized ORFs exhibit phenotypes, thereby providing clues as to the function of these genes (for example, *YBR077C*, *YDR101C*, *YIR016W* and *YIL129W*).

Interestingly, the exact site of transposon insertion within an ORF determines its disruption phenotype. As a result, our transposon can be used effectively to map functional domains within a given protein. For example, we have identified 17 transformants containing mTn insertions at 11 sites within *IMP2'*, a nuclear gene encoding a transcription factor involved in sugar utilization¹⁸. Transposon insertions within the amino terminus and central region of Imp2 (at amino acids 46, 69, 230 and 270) result in mutants unable to grow on glycerol, a non-fermentable carbon source. However, strains carrying an insertion at codon 334 (12 codons upstream of the translational stop codon) still grow normally on YPGlycerol. This same insertion, though, results in a defect in cell-wall synthesis (Fig. 3b) shared by strains containing insertions at residues 46, 230, 263 and 319. Therefore, nearly the entire length of Imp2 appears to be involved in cell-wall biogenesis¹⁷, although its carboxy terminus is not required for the control of sugar utilization. These findings illustrate the effectiveness of our collection in assigning specific functions to individual protein domains.

Many insertional mutants exhibit phenotypes under a specific subset of test conditions: these mutants appear to 'cluster' into discrete classes. To consider these clusters more rigorously, we have developed and implemented an algorithm by which haploid transformants may be grouped by common disruption phenotypes (see Methods). Based upon the tested phenotypes listed in Table 1, our collection of viable haploid transformants may be subdivided into 21 clusters: one cluster for each indicated growth condition as well as an additional cluster of mutants exhibiting wild-type phenotypes. Examples of these clusters are shown in Fig. 4. Transformants within a single cluster typically share a common mutant phenotype (Fig. 4) and can be subdivided further into smaller clusters. Interpreted judiciously, these phenotypic clusters provide a means of predicting cellular functions associated with a given ORF. For example, cluster 'YPG' contains transformants characterized most prominently by an inability to grow on YPGlycerol. This cluster is highly enriched in transformants carrying a transposon insertion within genes involved in cellular respiration (for example, *IFM1*, *IMP2'*, *SHY1* and *NDI1*). A number of strains within this cluster carry transposon insertions in previously uncharacterized genes (for example, *YLR294C*, *YJL131C* and *YGR101W*); from our clustering analysis, we can reasonably infer that these genes probably encode proteins involved in cellular respiration as well. Similar functional inferences may be drawn from each remaining data cluster (accessible in full at <http://bioinfo.mbb.yale.edu/genome/phenotypes>).

Additionally, our algorithm may be used to 'double cluster'

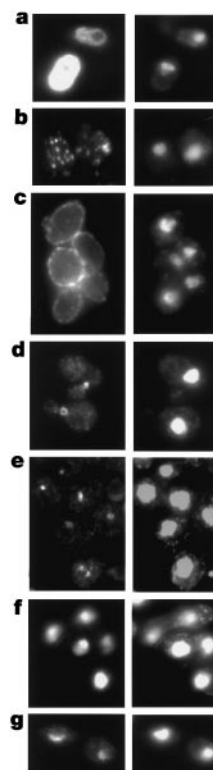


Figure 5 Immunolocalization of epitope-tagged proteins. Left, examples of immunofluorescence patterns in vegetative cells stained with monoclonal antibody against HA. Right, the same cells stained with the DNA-binding dye 4',6-diamidino-2-phenylindole (DAPI). **a**, Diffuse cytoplasmic staining in a transformant HAT-tagged through an in-frame mTn insertion event in *YLR249W*. **b**, A punctate pattern of cytoplasmic staining resulting from HAT-tagging of Crn1, an actin-binding protein that bundles actin filaments²⁹. **c**, Localization to the plasma membrane. **d**, Localization of HAT-tagged Bni4 to a ring in the mother-bud neck. **e**, Staining of the spindle body in a strain containing a HAT-tagged allele of the spindle pole protein Spc105. **f**, Nuclear staining observed in transformants HAT-tagged by mTn insertion in *NSR1*, a gene whose protein product binds nuclear localization sequences³⁰. **g**, Localization of HAT-tagged Nan1 to the nucleolus.

observed phenotypic data: we can use this algorithm to group phenotypic assays (rather than transformants) based upon the ability of these assays to identify strains exhibiting similar phenotypic profiles. By double-clustering the data presented above, we can identify sets of assays that are effective in selecting functionally related proteins with a greater degree of confidence. To illustrate this, consider the clustered assays presented in the dendrogram to the left of Table 1 (reproduced at the top of Fig. 4). Note the tight grouping of assays (for example, hypersensitivity to calcofluor, hygromycin and SDS) that are expected to identify mutants defective in cell-wall biogenesis and maintenance¹⁷. Similarly, assays for growth in hydroxyurea and MMS (methyl methanesulphonate) are closely associated; these assays would be expected to preferentially identify mutants defective in DNA metabolism. The use of such assays in combination would, therefore, constitute a more effective means of identifying a pool of mutants enriched in a common phenotype.

Further indication of cellular function may be inferred from the subcellular localization of a given protein. To determine these localization patterns, we have examined 1,340 diploid strains carrying in-frame HAT tag insertions by indirect immunofluorescence with antibodies directed against the HA-epitope. Localization of the HAT-tagged protein to a discrete cellular site was observed and confirmed for 201 strains; 214 additional strains showed cytoplasmic localization. A variety of localization patterns was detected within these strains: tagged proteins were localized to the nucleus, nucleolus, mitochondria, plasma membrane, cell neck and

spindle pole body (Table 2 and Fig. 5). In many cases, strains that show a distinct localization pattern contain a HAT tag in a known protein whose localization had previously been determined (for example, Top1–HAT¹⁹ and Rap1–HAT²⁰ localize to the nucleus). In other cases, however, our data constitute information concerning proteins that have not been studied or localized previously (such as the YHR196W gene product—a protein of unknown function—which localizes to the nucleolus).

Our transposon-based system also presents a means of defining gene expression profiles on a genomic scale. For example, to identify genes whose expression is strongly induced during sporulation, diploid transformants carrying mTn insertions were incubated on sporulation and control media and then assayed for β -gal activity. Of 106 insertions in yeast genomic DNA, 36 represent in-frame fusions in 31 different genes. Six of these genes have been previously shown to be induced in meiosis, eight are in named genes not known to be meiotically induced, and the remainder are in ORFs not yet characterized (Fig. 2). As noted above for fusion genes expressed during vegetative growth, mTn insertions are also in the antisense direction or in the +1 or –1 reading frame of annotated ORFs. In addition, several insertions are in NORFs, noncoding sequences or the ribosomal DNA.

In ref. 21, a DNA microarray representing all annotated ORFs in the yeast genome was screened for changes in gene expression during sporulation. Of the 31 meiotic genes we identified by in-frame *lacZ* fusions, only 17 were also found to be induced at least two-fold by microarray screening. For the remainder, microarray analysis indicated little or no induction during meiosis. One of the genes detected in our analysis but not by microarray screening is *MER1*, which has been shown to be strongly induced in meiosis both by northern blot hybridization and analysis of a *Mer1*– β -gal translational fusion gene²². This observation indicates that the β -gal assay may serve as a more sensitive indicator of meiotic gene expression and/or detect differences in protein levels more readily. In any case, our transposon-based approach is an effective means of identifying novel sporulation-induced genes; one gene identified in our screen is *YIL073C*, which we found to be important for spore viability and synaptonemal complex formation.

Our transposon insertion strategy allows the rapid generation of reporter gene fusions, epitope-tagging constructs and insertion alleles (both large and small), all from a single mutagenic event. As we do not bias this event towards a population of previously annotated genes, our approach effectively identifies new genes. This random approach, however, presents an obstacle in achieving saturation mutagenesis: small genes are less likely to be mutagenized by random mTn insertion than are large genes. Using our current approach, an additional 30,000 mTn insertions in yeast ORFs would be required to mutagenize 90% of the yeast genome. Once complete, however, we see our collection ushering in the 'new yeast genetics'—a new paradigm by which yeast genetic screens may be performed. Using a collection such as ours, researchers will be able to screen large numbers of mutant yeast strains for a given phenotype (by macroarray analysis or other methods) and instantly know the gene affected in a mutant of interest. Furthermore, the availability of alleles as a plasmid collection will allow introduction into any desired yeast strain background before screening. We have already distributed over 700 insertion alleles to the research community. These alleles will expedite gene analysis. □

Methods

Generation and analysis of yeast transformants

We used transposon mutagenesis in *E. coli* to generate $>10^6$ independent transformants; individual transformant colonies were selected and stored in 96-well plates¹⁰. Plasmids were prepared from these strains by alkaline lysis²³, digested with *NotI* and transformed into diploid strain Y800 (ref. 8). Following transformation, yeast cells were cultured as described¹⁰; after incubation under appropriate growth conditions, cells were permeabilized and assayed for β -gal activity¹⁰. Transformants displaying activity were arrayed in a 96-well format; each strain has a unique coordinate in this collection. The strains were

subsequently re-tested for β -gal activity after vegetative growth and sporulation. The results of this retest are accessible at <http://ycmi.med.yale.edu/ygac/triples.htm>.

To determine the genomic site of mTn insertion within these clones, plasmid-borne insertion alleles were sequenced using a primer complementary to bases 74–97 of mTn–3xHA/*lacZ* (GenBank accession number U54828). The resulting sequence data were searched against the *S. cerevisiae* genome using BLAST²⁴.

Phenotypic macroarray analysis

DNA obtained from the insertion allele plasmid collection was digested with *NotI* and transformed into haploid strain Y2279 [*MATa ura3-52 trp1 Δ ade2-101 lys2-801 leu2 Δ 98*]. Transformants were selected on SC-ura. Because phenotypes are often associated with mutations unlinked to the transposon insertion (M.S., unpublished data), we usually analysed two independent transformants. We scored yeast strains as exhibiting a defect only if both transformants displayed the defect or if two independent insertion alleles in the same gene produced the same defect. In many cases, haploid transformants were not obtained. We presume that in these cases either the affected gene was essential, or there was little transforming DNA. Transformants were grown in SC-ura medium overnight at 30 °C. A few microlitres of cell suspension were then transferred to selective growth medium using a 96-pin device for phenotypic macroarray analysis.

Cluster analysis

A combination of *k*-means clustering²⁵ and a heuristic algorithm was used to cluster the phenotypic macroarray data. The data were arranged into a 7,847 by 20 matrix *X*, with each row representing one transformant and each column representing one growth condition. The qualitative phenotype observations 'wild-type', 'weak', 'medium' and 'strong' were mapped onto the numerical scores 1, 3, 9 and 27, which reflect the different weights of the observations. The matrix rows were clustered based on the Euclidean distance between them, which is defined as $d_{ij} = \sqrt{\sum_{k=1}^N (x_{ik} - x_{jk})^2}$ where x_{ik} and x_{jk} are the elements in rows *i* and *j* and column *k* of matrix *X* and *N* = 20. As initial cluster centroids we used 20 data rows that exhibit the phenotype 'strong' in only one of the growth conditions plus one data row with no 'strong' phenotype. The clustering was performed in several iterations until the cluster centroids remained unchanged.

The tree describing the relationship between growth conditions was calculated with a program in the PHYLIP package²⁶ that performs the Fitch–Margoliash least-squares algorithm²⁷ on a 20 by 20 matrix describing the Euclidean distance between the columns of *X*. The Euclidean distance between columns is defined as in the formula above with *N* = 7,847, *i* and *j* representing columns, and *k* representing a row.

Production and immunolocalization of HAT-tagged strains

HAT tags were generated *in vivo* by *cre-lox* site-specific recombination¹⁰. HAT-tagged strains were analysed by indirect immunofluorescence as described^{8,10}, except that the primary antibody was mouse monoclonal anti-HA 16B12 (MMS101R, BAbCo) and the secondary antibody was Cy3-conjugated goat anti-mouse IgG (Jackson Labs).

Received 2 August; accepted 7 October 1999.

- Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 546–567 (1996).
- Mewes, H. W. *et al.* Overview of the yeast genome. *Nature* **387**, 7–65 (1997).
- Hieter, P. & Boguski, M. Functional genomics: it's all how you read it. *Science* **278**, 601–602 (1997).
- Velculescu, V. E. *et al.* Characterization of the yeast transcriptome. *Cell* **88**, 243–251 (1997).
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680–686 (1997).
- Shoemaker, D. D., Lashkari, D. A., Morris, D., Mittmann, M. & Davis, R. W. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nature Genet.* **14**, 450–456 (1996).
- Smith, V., Chou, K. N., Lashkari, D., Botstein, D. & Brown, P. O. Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science* **274**, 2069–2074 (1996).
- Burns, N. *et al.* Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* **8**, 1087–1105 (1994).
- Ross-Macdonald, P. R., Sheehan, A., Roeder, G. S. & Snyder, M. A multipurpose transposon system for analyzing protein production, localization, and function in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **94**, 190–195 (1997).
- Ross-Macdonald, P. *et al.* Methods for large-scale analysis of gene expression, protein localization, and disruption phenotypes in *Saccharomyces cerevisiae*. *Methods Mol. Cell. Biol.* **5**, 298–308 (1995).
- Cherry, M. *et al.* SGD: *Saccharomyces* genome database. *Nucleic Acids Res.* **26**, 73–79 (1998).
- Schatz, P. J., Pillus, L., Grisafi, P., Solomon, F. & Botstein, D. Two functional alpha-tubulin genes of the yeast *Saccharomyces cerevisiae* encode divergent proteins. *Mol. Cell. Biol.* **6**, 3711–3721 (1986).
- Page, B. D. & Snyder, M. Clk1: a developmentally regulated spindle pole body-associated protein important for microtubule functions in *Saccharomyces cerevisiae*. *Genes Dev.* **6**, 1414–1429 (1992).
- Stearns, T., Hoyt, M. A. & Botstein, D. Yeast mutants sensitive to antimicrotubule drugs define three genes that affect microtubule function. *Genetics* **124**, 251–262 (1990).
- Evangelista, C. C. Jr, Rodriguez, T. A. M., Limbach, M. P. & Zitomer, R. S. Rox3 and Rts1 function in the global stress response pathway in baker's yeast. *Genetics* **142**, 1083–1093 (1996).
- White, W. H. & Johnson, D. I. Characterization of synthetic-lethal mutants reveals a role for the *Saccharomyces cerevisiae* guanine-nucleotide exchange factor Cdc24p in vacuole function and Na⁺ tolerance. *Genetics* **147**, 43–55 (1997).
- Lussier, M. *et al.* Large scale identification of genes involved in cell surface biosynthesis and architecture in *Saccharomyces cerevisiae*. *Genetics* **147**, 435–450 (1997).
- Lodi, T., Goffrini, P., Ferrero, I. & Donnini, C. IMP2, a gene involved in the expression of glucose-repressible genes in *Saccharomyces cerevisiae*. *Microbiology* **141**, 2201–2209 (1995).
- Christman, M. F., Dietrich, F. S. & Fink, G. R. Mitotic recombination in the rDNA of *S. cerevisiae* is suppressed by the combined action of DNA topoisomerases I and II. *Cell* **55**, 413–425 (1988).
- Shore, D. & Nasmyth, K. Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* **51**, 721–732 (1987).



Microtubule dynamics
Temperature sensitivity
Sporulation-induced
Osmo- and cation sensitivity
No phenotype, untested

Multiple phenotypes
Respiration
Haploid invasive mutants
DNA metabolism and protein synthesis
Cell wall biogenesis/maintenance

21. Chu, S. *et al.* The transcriptional program of sporulation in budding yeast. *Science* **282**, 699–705 (1998).
22. Engebrecht, J. & Roeder, G. S. Mer1, a yeast gene required for chromosome pairing and genetic recombination, is induced in meiosis. *Mol. Cell. Biol.* **10**, 2379–2389 (1990).
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Press, Cold Spring Harbor, New York, 1989).
24. Altschul, S. F., Gish, W., Miller, W., Meyers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
25. Kaufman, L. & Rousseeuw, P. J. *Finding Groups in Data* (Wiley, New York, 1990).
26. Felsenstein, J. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166 (1989).
27. Fitch, W. M. & Margoliash, E. Construction of phylogenetic trees. *Science* **155**, 279–284 (1967).
28. Adams, A., Gottschling, D. E., Kaiser, C. A. & Stearns, T. *Methods in Yeast Genetics* (Cold Spring Harbor Press, Cold Spring Harbor, New York, 1997).
29. Goode, B. L. *et al.* Coronin promotes the rapid assembly and cross-linking of actin filaments and may link the actin and microtubule cytoskeletons in yeast. *J. Cell Biol.* **144**, 83–89 (1999).
30. Lee, W. C., Xue, Z. X. & Melese, T. The NSR1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs. *J. Cell Biol.* **113**, 1–12 (1991).

Acknowledgements

P.S.R.C. is supported by the Fundacao de Amparo a Pesquisa do Estado de Sao Paulo, Brazil. This work was supported by an NIH grant (to G.S.R. and M.S.).

Correspondence and requests for materials should be addressed to M.S. (e-mail: michael.snyder@yale.edu).

Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast

John J. Wyrick^{*†}, Frank C. P. Holstege^{*}, Ezra G. Jennings^{*†}, Helen C. Causton^{*}, David Shore[‡], Michael Grunstein[§], Eric S. Lander^{*†} & Richard A. Young[†]

^{*} Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

[†] Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[‡] Department of Molecular Biology, University of Geneva, Geneva 4, CH-1211, Switzerland

[§] Department of Biological Chemistry, UCLA School of Medicine and Molecular Biology Institute, University of California, Los Angeles, California 90095, USA

Eukaryotic genomes are packaged into nucleosomes, which are thought to repress gene expression generally^{1–3}. Repression is particularly evident at yeast telomeres, where genes within the telomeric heterochromatin appear to be silenced by the histone-binding silent information regulator (SIR) complex (Sir2, Sir3, Sir4) and Rap1 (refs 4–10). Here, to investigate how nucleosomes and silencing factors influence global gene expression, we use high-density arrays to study the effects of depleting nucleosomal histones and silencing factors in yeast. Reducing nucleosome content by depleting histone H4 caused increased expression of 15% of genes and reduced expression of 10% of genes, but it had little effect on expression of the majority (75%) of yeast genes. Telomere-proximal genes were found to be de-repressed over regions extending 20 kilobases from the telomeres, well beyond the extent of Sir protein binding^{11,12} and the effects of loss of Sir function. These results indicate that histones make Sir-independent contributions to telomeric silencing, and that the role of histones located elsewhere in chromosomes is gene specific rather than generally repressive.

We used a yeast strain (*Saccharomyces cerevisiae* UKY403) in which the sole copy of histone H4 is under the control of the *GAL1* promoter^{13–15} to examine the effects of nucleosome depletion on expression of all protein-coding genes. Nucleosome depletion begins about one hour after switching from galactose to glucose

medium, leading to a reduction in the nucleosomal content of yeast cells by around 50% within six hours^{14,15}. Duplicate cultures of strain UKY403 and a control strain with a wild-type histone H4 gene were switched to glucose medium and harvested over a time course and labelled antisense RNA (cRNA) prepared from purified messenger RNA was hybridized to Affymetrix GeneChip arrays¹⁶. Interactive databases supporting this work, details of the experimental reagents and approaches and additional information can be obtained on the World-Wide Web at <http://web.wi.mit.edu/young/chromatin/>. For a summary of the data sets described here, see Supplementary Information.

The results of the time course of histone depletion are summarized in Fig. 1. Of the 5,894 genes whose mRNA levels were detected at the 6-h time point, the expression of 75% was not significantly altered, the expression of 15% increased more than three-fold, and the expression of 10% decreased more than three-fold during histone H4 depletion. It was notable that expression of most yeast genes was unaffected by histone depletion. Previous studies have shown that nucleosome loss in histone H4 depletion experiments results in a general reduction in nucleosome density¹⁵, and this reduction occurs at all individual genes tested^{14,17,18}, so it is likely that most genes experience a significant reduction in nucleosome density. Reduced nucleosome density may not affect expression of genes that are already induced; inducible genes have been found to be sensitive to nucleosomes loss only when in their uninduced state^{13,14}. It is also possible that nucleosomes have minimal roles in regulation of many genes; transcriptional activators and repressors may be dominant in regulation of such genes, as in prokaryotic cells.

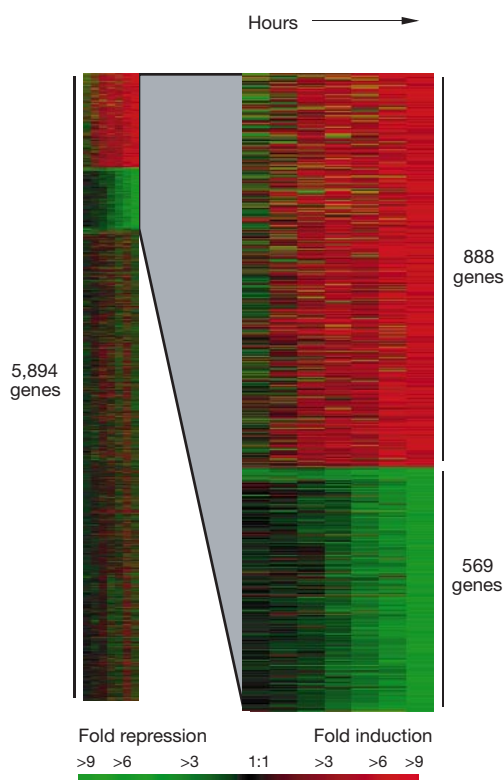


Figure 1 Effect of histone H4 depletion on genome-wide expression. The data for each gene are displayed as the ratio of mRNA in histone H4-depleted cells to that of wild-type cells harvested 0, 0.5, 1, 1.5, 2, 4 and 6 h after a shift from galactose to glucose medium. Each horizontal strip represents a single gene. The fold change is represented by a colour (see colour scale). The left panel shows genes whose mRNA levels increase (top 15% of diagram), genes whose mRNA levels decrease (next 10%), and genes which were relatively unaffected over the time course (bottom 75% of diagram). In the panel on the right, the subset of genes whose mRNA levels increased or decreased more than 3-fold at the 6-h time point are displayed. See ref. 30 for the display program.

A substantial number of genes exhibited decreased expression levels during histone depletion, and only a fraction of these could be attributed to indirect effects, indicating that nucleosomes may positively regulate expression of many genes. To determine how the slow growth and cell-cycle arrest phenotypes of histone-depleted cells may have contributed to changes in gene expression, comparisons were made with the expression profiles of cells whose growth slows during the diauxic shift¹⁹ or which are arrested in the G2/M phase of the cell-cycle by nocadazole (see Supplementary Information). Of the genes that show decreased expression in the histone H4 depletion experiment, only half could be so affected as a consequence of slow growth or cell-cycle arrest. Although histones and nucleosomes are known to positively regulate the expression of some specific genes, including the mouse mammary tumour virus (MMTV), *Xenopus* vitellogenin B1 and *Tetrahymena* CyP genes^{18,20,21}, our results indicate that positive roles for nucleosomal histones in gene expression are more pervasive than previously suspected.

When we examined the distribution of effects on gene expression across yeast chromosomes, we found clusters of de-repressed genes at the telomeres. The effects on several representative chromosomes are shown in Fig. 2a and the effects on all chromosomes can be seen at the authors' Web site. About 50% of genes within 20 kilobases (kb) of telomeres showed increased mRNA levels during the histone depletion experiment, contrasting with the 15% of genes located elsewhere in the genome that exhibited increased levels. Figure 3 shows the change in transcript levels of genes within 20 kb of all 32 yeast telomeres over the time course of histone H4 depletion. De-repression of telomere-proximal genes occurred at essentially all telomeres and began early in the time course, indicating that de-repression of these genes is likely to be a direct consequence of histone loss.

To define the extent of the clustering of repressed genes near telomeres, we examined how the fraction of genes that are de-repressed upon histone H4 depletion varied with distance from the

Table 1 Summary of telomere-proximal genes affected by histone H4 depletion and *SIR3* deletion

Distance from telomere	Histone H4 depletion		<i>SIR3</i> deletion	
	Fraction of genes de-repressed	χ^2	Fraction of genes de-repressed	χ^2
0–10 kb	56%	70	16%	74
10–16 kb	52%	57	2%	0.1
16–21 kb	28%	8	2%	0.1
21–24 kb	26%	5	0%	0.7
25–28 kb	26%	5	0%	0.7
28–32 kb	22%	2	2%	0.1
32–36 kb	28%	8	0%	0.7
36–39 kb	8%	2	0%	0.7
39–42 kb	16%	0.1	4%	2
Genome	15%	–	1.5%	–

Genes closest to a telomere were examined in consecutive intervals of 50 genes, and the fraction of genes de-repressed in each mutant was calculated for each interval. A χ^2 value for each interval was calculated by comparing the fraction of genes de-repressed in the interval with the genome-wide average.

telomere (Fig. 4a). To test the statistical significance of the clustering of repressed genes near telomeres, we compared the observed fraction of telomeric genes that were de-repressed by histone H4 depletion with the genome-wide average using a χ^2 test. The results (see Table 1) show that the proportion of genes repressed by histones is statistically significantly higher within 18–20 kb of the chromosome ends than in general across the genome. Furthermore, examination of genome-wide expression data from wild-type cells²² revealed that genes within 20 kb of a telomere have on average 0.5 mRNA molecules per cell, five times less than the average number for the entire genome (2.4 molecules per cell). Hence, genes within 20 kb of telomeres are silenced in wild-type cells, and this repression is dependent on wild-type levels of nucleosomes.

The proteins Sir2, Sir3, Sir4 and Rap1 are thought to be required for telomeric silencing, with Sir3 being particularly important^{4,6,11,23}. However, when the effects of deleting *SIR2*, *SIR3* or *SIR4* on genome-wide expression were investigated, only 7–9% of the

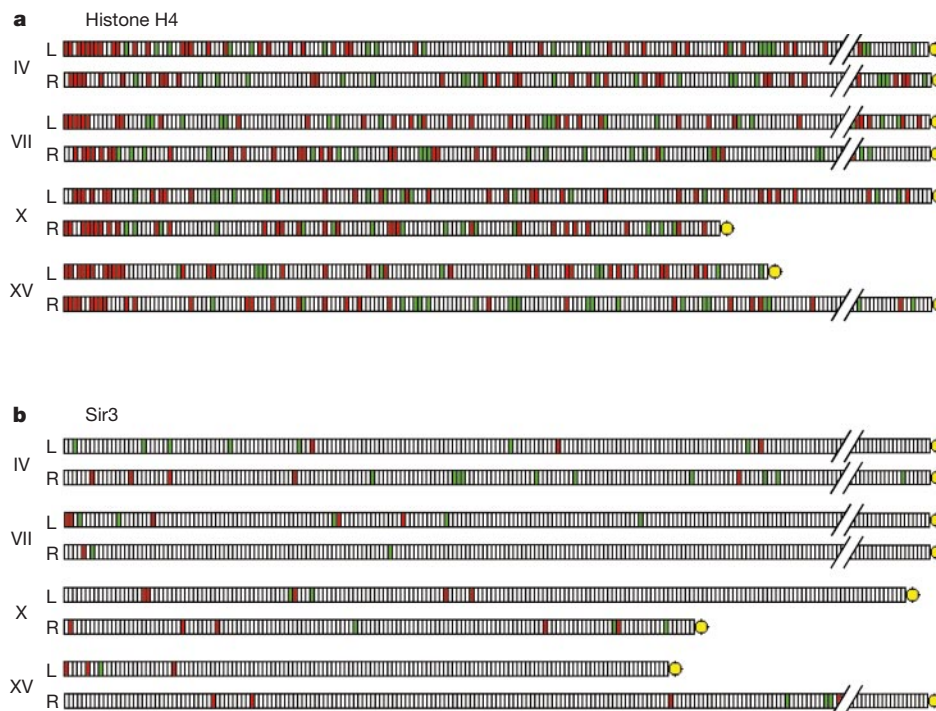


Figure 2 Chromosomal display of changes in gene expression. **a**, Histone H4 depletion (6-h time point); **b**, deletion of *SIR3*. Each gene is represented by a rectangle, the colour of which indicates whether there has been >3-fold increase in its mRNA (red) or >3-fold decrease in its mRNA (green) relative to control. The genes are ordered according to their

normal positions along each chromosome in half-chromosome segments, beginning at the telomere of the chromosome (left or right) and ending at the centromere (yellow circle). Selected chromosomes are displayed; all chromosomes can be viewed at web.wi.mit.edu/young/chromatin/.

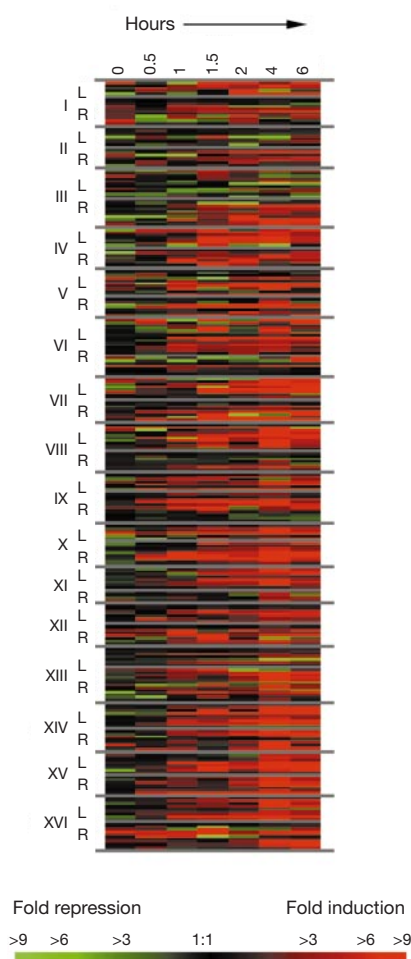


Figure 3 Effect of histone H4 depletion on telomere-proximal gene expression. Genes within 20 kb of a chromosome end were selected and displayed as in Fig. 1. Genes are grouped and labelled by chromosome and telomere, and displayed in chromosomal order, with the top gene in each set closest to the end of the chromosome.

genes within 20 kb of telomeres were found to be de-repressed; the results for the *SIR3* mutant are shown in Fig. 2b. As expected^{4,24}, expression of genes known to be repressed by the $\alpha 1/\alpha 2$ repressor decreased in *SIR2*, *SIR3* or *SIR4* mutants, and expression of the *URA3* reporter gene integrated at telomere VII-L in the *SIR2* mutant was increased seven-fold. Furthermore, the loss of these silencing factors clearly leads to de-repression of genes located very close to telomere repeats. Finally, mating-type-regulated genes are not responsible for the telomeric expression profile of *SIR* mutants, as $\alpha 1$ -deficient *SIR4* mutant haploid cells exhibited the same pattern of telomeric gene induction as $\alpha 1$ -containing *SIR4* mutant haploids (see authors' Web site).

Loss of Sir function appears to have a major effect on expression of genes within 6–8 kb of telomeres, whereas the effects of histone H4 depletion appear to extend to nearly 20 kb (Fig. 4a; Table 1). In other words, the fraction of genes that are de-repressed is reduced with distance from the telomere in both mutants, but the decline appears to be much sharper for *SIR3* than for histone H4. We tested this observation statistically by using a maximum likelihood approach to test the hypothesis that the rate of decrease in the two mutants is the same (for details, see authors' Web site). The results indicate that the rate of decrease in the two mutants is distinct ($P < 0.02$), with the effects in the *SIR3* mutant dropping off more quickly as a function of distance from the telomere than in the histone H4 depletion mutant. Figure 4b summarizes the effects of histone and *SIR3* loss on expression of telomere-proximal

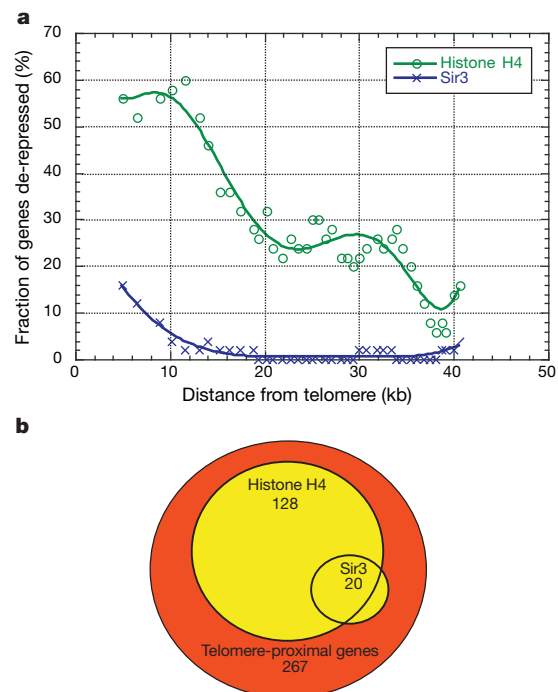


Figure 4 Extent of telomeric silencing by histone H4 and Sir3. **a**, To define the extent of the clustering of repressed genes near telomeres, we examined how the proportion of genes that were de-repressed upon histone H4 depletion or *SIR3* deletion varied with distance from the telomere. The fraction of genes that were de-repressed and average distance from telomere for each sliding 50-gene window were plotted for both mutants. **b**, Venn diagram showing how histone H4 depletion and *SIR3* deletion affects telomere-proximal gene expression. The diagram depicts all telomere-proximal genes (large red circle), and the subset of these genes that are de-repressed by histone H4 depletion (large yellow circle) or *SIR3* deletion (small yellow circle). A table containing a list of all 267 telomere-proximal genes, and the effects of histone H4 depletion and deletion mutants of *SIR2*, *SIR3*, *SIR4* and *RAP1* on gene expression, is available at web.wi.mit.edu/young/chromatin/.

genes, and shows that the set of telomere-proximal genes that is de-repressed by *SIR3* loss is predominantly a subset of those de-repressed by histone H4 depletion.

These data support the current model of telomeric heterochromatin in which Sir proteins are required for the silencing of telomere-proximal genes located very close to the ends of chromosomes. However, telomere-proximal DNA was repressed over regions extending 20 kb from telomeres, well beyond the 3–4 kb seen for Sir protein binding^{11,12} and the effects due to loss of Sir function, indicating that the repression of subtelomeric genes is accomplished by a mechanism distinct from that used at telomeric heterochromatin. Genome-wide expression analysis of a Rap1 silencing-defective mutant (*rap1-17*) showed that its contribution to telomeric silencing is similar to that of Sir proteins (see Supplementary Information or authors' Web site). A search of expression databases for additional factors that might contribute to repression of subtelomeric genes revealed that deletion of *TUP1* (ref. 19) de-represses around 13% of genes with 20 kb of telomeres, thus accounting for a small fraction of the genes repressed by histones, but not by Sir proteins. It will be instructive to obtain genome-wide expression data for additional proteins that may contribute to telomeric silencing, including Cac1, which assembles histones and DNA into nucleosomes^{25,26}, yeast Ku proteins, which are involved in DNA repair and are required for the nuclear localization of telomeres²⁷, Hst1, 2, 3 and 4, which are homologues of Sir2 (ref. 28), and Set1 (ref. 29).

Our results show that histones make Sir-independent contribu-

tions to telomeric silencing, and indicate that the role of histones located elsewhere in chromosomes is gene specific rather than generally repressive. Knowledge of the effect of histone depletion on expression of all yeast genes provides the framework for future studies to identify the complete set of factors that are responsible for telomere-proximal silencing, and establishes the foundation for accurately reconstituting gene-specific transcription reactions *in vitro*. □

Methods

Strains, media and growth conditions

S. cerevisiae UKY403 (ref. 15; histone H4 depletion) and MHY308 (ref. 15; wild-type histone H4) cells were grown in parallel in galactose-containing medium (YEPGalactose) and switched to glucose-containing medium (YPD). Cells were grown for various times (0, 0.5, 1, 1.5, 2, 4, 6 h) in YPD before harvesting. The final OD₆₀₀ for all time points was 0.5–0.7. All other strains were grown in YPD media until reaching OD₆₀₀ = 0.5–0.7 (see authors' Web site for more details and a complete list of all strains used).

Genome-wide expression profiling

Biotin-labelled cRNA target was prepared as described²². The cRNA was hybridized to a set of four oligonucleotide arrays (GeneChip Ye6100 arrays, Affymetrix) and scanned as described¹⁶. Intensities were captured using GeneChip software (Affymetrix) and a single raw expression level for each gene was determined.

Data analysis

The raw data from each chip were normalized and corrected as described²². A change in mRNA levels was deemed significant and reported in the lists of genes whose expression was increased or decreased more than 3-fold based on the following criteria: the average fold change was more than 3-fold, the fold change in each experiment was greater than 1.5-fold and the change in the values was above background values in both comparisons. See authors' Web site for more details.

Telomere-proximal gene analysis and statistics

Genes within 40 kb of a telomere throughout the genome were pooled and ordered according to their distance from a telomere. Genes with sequence homologues whose mRNA levels could not be distinguished with certainty (genes with an Affymetrix 'f' call) were excluded from the analysis. The genes were grouped by their distance from a telomere either in consecutive intervals of 50 genes (Table 1) or in a sliding 50-gene window which was moved in 10-gene steps (Fig. 4a). For each interval, the fraction of genes that were depressed in each mutant was calculated. The change in the fraction of genes that were depressed with distance from telomere is shown in Fig. 4a. A χ^2 value for each interval was calculated by comparing the fraction of genes that were de-repressed in the interval to the genome-wide average (Table 1).

Received 16 July; accepted 6 September 1999.

- Paranjape, S. M., Kamakaka, R. T. & Kadonaga, J. T. Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Annu. Rev. Biochem.* **63**, 265–297 (1994).
- Workman, J. L. & Kingston, R. E. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **67**, 545–579 (1998).
- Gregory, P. D. & Horz, W. Life with nucleosomes: chromatin remodelling in gene regulation. *Curr. Opin. Cell Biol.* **10**, 339–345 (1998).
- Aparicio, O. M., Billington, B. L. & Gottschling, D. E. Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* **66**, 1279–1287 (1991).
- Kurtz, S. & Shore, D. RAP1 protein activates and silences transcription of mating-type genes in yeast. *Genes Dev.* **5**, 616–628 (1991).
- Kyrion, G., Liu, K., Kiu, C. & Lustig, A. J. RAP1 and telomere structure regulate telomere position effects in *Saccharomyces cerevisiae*. *Genes Dev.* **7**, 1146–1159 (1993).
- Loo, S. & Rine, J. Silencing and heritable domains of gene expression. *Annu. Rev. Cell. Dev. Biol.* **11**, 519–548 (1995).
- Zakian, V. A. Structure, function, and replication of *Saccharomyces cerevisiae* telomeres. *Annu. Rev. Genet.* **30**, 141–172 (1996).
- Grunstein, M. Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell* **93**, 325–328 (1998).
- Lustig, A. J. Mechanisms of silencing in *Saccharomyces cerevisiae*. *Curr. Opin. Genet. Dev.* **8**, 233–239 (1998).
- Hecht, A., Strahl-Bolsinger, S. & Grunstein, M. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* **383**, 92–96 (1996).
- Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* **11**, 83–93 (1997).
- Han, M. & Grunstein, M. Nucleosome loss activates yeast downstream promoters *in vivo*. *Cell* **55**, 1137–1145 (1988).
- Han, M., Kim, U. J., Kayne, P. & Grunstein, M. Depletion of histone H4 and nucleosomes activates the PHO5 gene in *Saccharomyces cerevisiae*. *EMBO J.* **7**, 2221–2228 (1988).
- Kim, U. J., Han, M., Kayne, P. & Grunstein, M. Effects of histone H4 depletion on the cell cycle and transcription of *Saccharomyces cerevisiae*. *EMBO J.* **7**, 2211–2219 (1988).
- Wodicka, L., Dong, H., Mittmann, M., Ho, M. H. & Lockhart, D. J. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnol.* **15**, 1359–1367 (1997).
- Saunders, M. J., Yeh, E., Grunstein, M. & Bloom, K. Nucleosome depletion alters the chromatin structure of *Saccharomyces cerevisiae* centromeres. *Mol. Cell. Biol.* **10**, 5721–5727 (1990).

- Chavez, S. & Beato, M. Nucleosome-mediated synergism between transcription factors on the mouse mammary tumor virus promoter. *Proc. Natl Acad. Sci. USA* **94**, 2885–2890 (1997).
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. Exploring the metabolic and genetic control of gene expression of a genomic scale. *Science* **278**, 680–686 (1997).
- Schild, C., Claret, F. X., Wahli, W. & Wolffe, A. P. A nucleosome-dependent static loop potentiates estrogen-regulated transcription from the *Xenopus* vitellogenin B1 promoter *in vitro*. *EMBO J.* **12**, 423–433 (1993).
- Shen, X. & Gorovsky, M. A. Linker histone H1 regulates specific gene expression but not global transcription *in vivo*. *Cell* **86**, 475–483 (1996).
- Holstege, F. C. *et al.* Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
- Moretti, P., Freeman, K., Coddly, L. & Shore, D. Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* **8**, 2257–2269 (1994).
- Rine, J. & Herskowitz, I. Four genes responsible for a position effect on expression from HML and HMR in *Saccharomyces cerevisiae*. *Genetics* **116**, 9–22 (1987).
- Enomoto, S., McCune-Zierath, P. D., Gerami-Nejad, M., Sanders, M. A. & Berman, J. RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function *in vivo*. *Genes Dev.* **11**, 358–370 (1997).
- Kaufman, P. D., Kobayashi, R. & Stillman, B. Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. *Genes Dev.* **11**, 345–357 (1997).
- Laroche, T. *et al.* Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr. Biol.* **8**, 653–656 (1998).
- Brachmann, C. B. *et al.* The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. *Genes Dev.* **9**, 2888–2902 (1995).
- Nislow, C., Ray, E. & Pillus, L. SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* **8**, 2421–2436 (1997).
- Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14863–14969 (1998).

Supplementary Information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank T. Golub, D. Gottschling, C. Hengartner, T. Lee, H. Madhani and C. Wilson for helpful discussions, and N. Hannett, M. Gaasenbeek and C. Huard for technical support. This work was supported by funds from the NIH, Bristol-Myers Squibb Company, Affymetrix Inc., Millennium Pharmaceuticals Inc., and the Whitehead Institute Leadership Circle. J.J.W. is a predoctoral fellow of the National Science Foundation, F.C.P.H. was supported by fellowships from EMBO and the Hujan Frontier Science Program and E.G.J. is a predoctoral fellow of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to R.A.Y. (e-mail: young@wi.mit.edu).

LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite

N. Toni*, P.-A. Buchs*, I. Nikonenko*, C. R. Brion† & D. Muller*

* Neuropharmacology, CMU, University of Geneva, 1211 Geneva 4, Switzerland
† Institute of Anatomy, University of Bern, 3000 Bern, Switzerland

Structural remodelling of synapses^{1–4} and formation of new synaptic contacts^{5–8} has been postulated as a possible mechanism underlying the late phase of long-term potentiation (LTP), a form of plasticity which is involved in learning and memory⁹. Here we use electron microscopy to analyse the morphology of synapses activated by high-frequency stimulation and identified by accumulated calcium in dendritic spines. LTP induction resulted in a sequence of morphological changes consisting of a transient remodelling of the postsynaptic membrane followed by a marked increase in the proportion of axon terminals contacting two or more dendritic spines. Three-dimensional reconstruction revealed that these spines arose from the same dendrite. As pharmacological blockade of LTP prevented these morphological changes, we conclude that LTP is associated with the formation of new, mature and probably functional synapses contacting the same presynaptic terminal and thereby duplicating activated synapses.

Electron microscopic analyses of the morphological changes associated with LTP are difficult without being able to identify activated synapses¹⁰. Here we used a potassium chromium–tris-oxalate precipitation protocol during osmium fixation¹¹ to reveal, as a fine precipitate, the calcium that had accumulated in dendritic spines following high-frequency stimulation (Fig. 1a). Previous work had shown that this method is specific for calcium, that the precipitate is only present in a small proportion of spine profiles under control conditions, and that theta burst stimulation (TBS) to a group of CA3 neurons results in an NMDA (*N*-methyl-D-aspartate)-dependent increase in the number of CA1 spine profiles containing calcium precipitate⁴ (Fig. 1b). Analysis of the density of precipitate in spine profiles (Fig. 1c) shows a bimodal distribution with a clear distinction between labelled (black columns) and unlabelled (grey columns) profiles. The small number of profiles for which classification was unclear (white column, 1–2% of profiles) were discarded from quantitative analyses.

Using this approach, we examined the ultrastructural correlate of LTP as a function of time after induction of synaptic potentiation. LTP was induced at a group of synapses by application of TBS and their responses were monitored over various periods of time (5, 15, 30, 45, 60 and 120 min). A second TBS was applied 5 min before fixation to re-label the same synapses with calcium precipitate (Fig. 1d). The first notable change observed

in the population of labelled dendritic spines was an increase in the proportion of spine profiles having a partitioned postsynaptic density (PSD), also referred to as perforated synapses^{1–4,10,11} (Fig. 2a, b). The proportion of perforated synapses was $18.8 \pm 1\%$ in non-stimulated cultures; it increased gradually to $23.4 \pm 3.5\%$, $25 \pm 2.4\%$ and $45.8 \pm 2.8\%$ of all labelled synaptic profiles at 5, 15 and 30 min after stimulation, respectively ($n = 3–7$; 220–949 spine profiles analysed). Unexpectedly, the occurrence of these perforated synapses was only transient: their proportion decreased to $28.6 \pm 7.7\%$ after 45 min to reach control values after 60 min ($18.1 \pm 2.5\%$). Unbiased stereological measurements confirmed these changes ($15 \pm 2\%$, $38 \pm 2\%$ and $15 \pm 5\%$ at 5, 30 and 60 min, respectively; $n = 2$; 65–112 spine profiles). These changes were not detected when non-labelled synapses were analysed ($22.7 \pm 1.1\%$, $n = 6$, 246 spine profiles), and were accompanied by transitory changes in the perforation characteristics of the synapses. Figure 2e shows that the size of the perforation and of the spinule, characterized by a small protrusion of the postsynaptic membrane into the presynaptic terminal (Fig. 2f), was considerably greater 30 min after stimulation than at other time points (perforation size: 284 ± 29 nm at 30 min versus 157 ± 19 nm and 145 ± 14 nm at 0–15 min and 45–120 min, respectively, $n = 5–8$; spinule height: 167 ± 28 nm versus 62 ± 4 nm and 74 ± 5 nm; $n = 5–11$).

These transitory changes were followed, one hour after stimulation, by a threefold increase in the presence of multiple spine boutons (Fig. 2c, d). These were characterized by the presence of at least two dendritic spines contacting the same axon terminal. Under control conditions, when analysing non-stimulated slices,

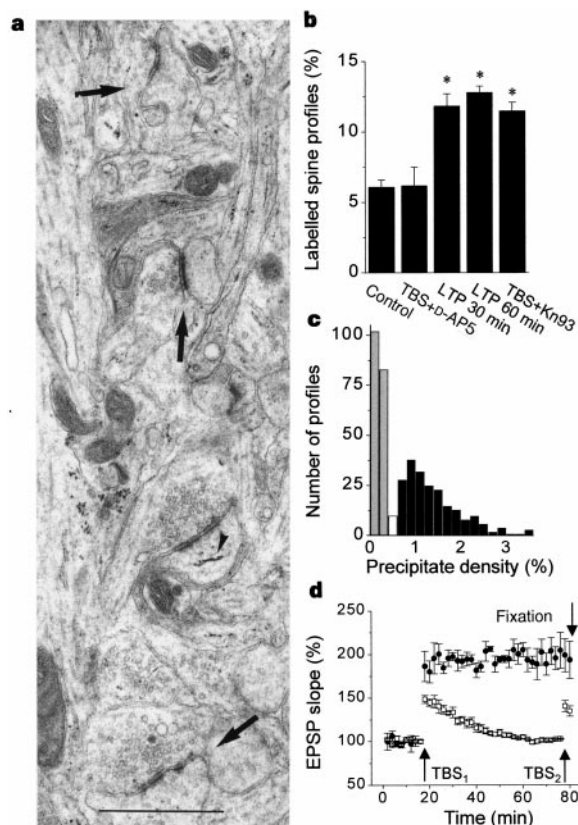


Figure 1 Labelling of activated spine profiles by induction of LTP. **a**, Low-magnification view of synaptic profiles in a slice culture fixed 5 min after high-frequency stimulation showing one labelled (arrowhead on labelling) and three unlabelled (arrows) spine profiles. Scale bar, 1 μ m. **b**, Proportion of spine profiles containing calcium precipitate under the various conditions tested (asterisk, $P < 0.001$). **c**, Distribution of the precipitate density, expressed in per cent and measured as the area of precipitate over the spine profile area, for the 60-min time point. Note the distinction between the populations of labelled (black columns) and unlabelled (grey columns) profiles. Profiles for which classification was unclear (white column) were discarded. **d**, Electrophysiological recordings of the LTP induced in control organotypic cultures (black circles) and following treatment with the kinase antagonist KN93 (open squares). After LTP induction (TBS₁), a second TBS was applied 5 min before fixation to re-label activated synapses.

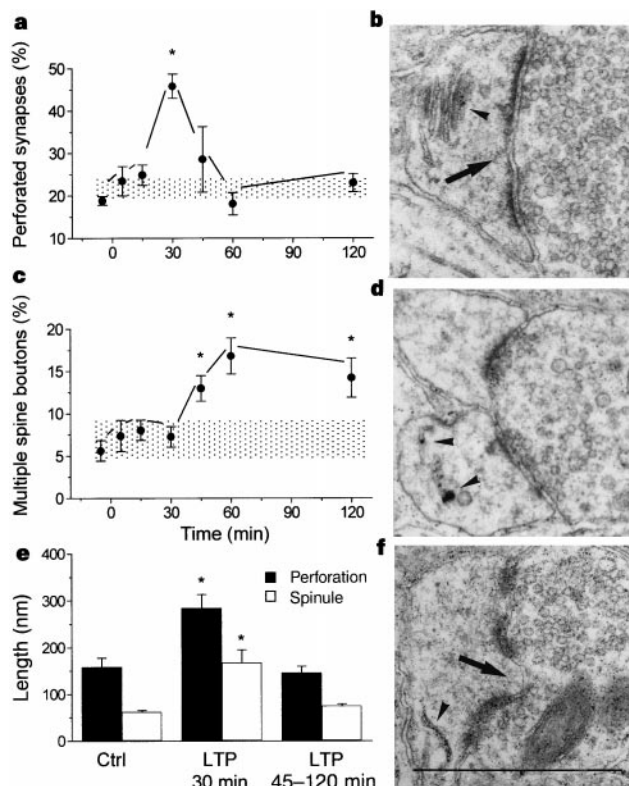


Figure 2 Time course of ultrastructural changes associated with LTP induction. **a**, Transient changes in the proportion of perforated synapses ($n = 3–7$; 220–630 spine profiles; asterisk, $P < 0.001$). **b**, A perforated spine profile (arrow, perforation). **c**, Increase in the proportion of multiple spine boutons 45–60 min after LTP induction ($n = 3–7$; 206–364 spine profiles; asterisk, $P < 0.01$). **d**, A multiple spine bouton. **e**, Morphometric changes in the size of the perforation and spinule after LTP induction ($n = 5–11$; 65–149 spine profiles; asterisk, $P < 0.01$). **f**, A perforated synapse exhibiting a prominent spinule (arrow). Calcium precipitates indicated by arrowheads. Scale bar, 1 μ m.

the proportion of multiple spine boutons was $5.6 \pm 1.2\%$ of labelled spine profiles. Similar values were found 5, 15 and 30 min after high frequency stimulation ($7.3 \pm 1.8\%$, $8 \pm 1.2\%$ and $7.3 \pm 1.2\%$, respectively; $n = 3-7$; 206–364 spine profiles). However, 45 min after LTP induction, the proportion of labelled, multiple spine boutons in which either one or both spine profiles contained the precipitate represented $13 \pm 1.5\%$ of all labelled spine profiles analysed. This value reached $16.8 \pm 2.1\%$ 1 h after LTP induction and $14.2 \pm 2.3\%$ after 2 h ($n = 3-4$, 223–308 spine profiles). No increase was detected when analysing non-labelled synapses ($5.9 \pm 0.7\%$ 60 min after LTP; $n = 3$, 323 spine profiles).

We used unbiased stereological methods to confirm these results. Analyses were made 5 and 60 min after LTP induction and showed the same increase in multiple spine boutons (5 min: $6.2 \pm 0.4\%$, $n = 2$, 64 spine profiles; 60 min: $15 \pm 2.5\%$, $n = 2$, 119 spine profiles). In addition, to establish further the link between LTP induction and these morphological changes, we carried out experiments in which LTP was pharmacologically blocked by treatment with the calcium/calmodulin protein kinase II antagonist KN93 (Fig. 1d). Labelling of spine profiles was not affected by KN93 treatment (Fig. 1b), but we did not detect any increase in perforated synapses or in multiple spine boutons 30 and 60 min after stimulation, respectively (Fig. 3). These data indicate that the morphological changes reported here are specifically associated with induction of LTP.

As the increase in multiple spine boutons indicated possible synaptogenesis, we used three-dimensional reconstruction of 75 multiple spine boutons to analyse the origins of the newly formed dendritic spines (Fig. 4). Analysis of unlabelled multiple spine boutons reconstructed 5, 45 and 60 min after stimulation revealed, in agreement with previous studies¹², that in $89 \pm 8\%$ of cases the two spines arose from different dendrites (mean $\pm$ s.e.m. of six cultures, but 38 out of 41 reconstructed boutons; Fig. 4a, b). A similar result was found when analysing labelled multiple spine boutons reconstructed from a slice culture fixed 5 min after stimulation (spines arose from different dendrites in eight out of nine cases). In contrast, however, reconstruction of 25 labelled multiple spine boutons from four cultures fixed 45–60 min after LTP induction revealed that in 15 of them the two spines arose from the same dendrite. The mean proportion of multiple spine boutons in which

the two spines arose from the same dendrite was thus $66 \pm 14\%$ in four cultures as compared to $11 \pm 8\%$ for unlabelled multiple spine boutons ($n = 6$; $P < 0.01$). As shown in Fig. 4a, the increase in multiple spine boutons measured on single sections essentially resulted from cases in which the two spines arose from the same dendrite (Fig. 4c–e). This new population of duplicated synapses thus clearly indicated that new synapses were being formed between the same axon terminal and dendrite.

The morphology of these newly formed spines was similar to that of mature dendritic spines (Fig. 4e). They had well defined necks and spine heads, sometimes including a spine apparatus, and a clearly visible PSD facing a pool of presynaptic vesicles. Furthermore, for the labelled multiple spine boutons observed after 45–60 min and contacting spines originating from the same dendrite, serial section analysis indicated that in 9 out of 15 cases calcium precipitate was clearly visible in both spines. Such cases were also frequently observed on single section analyses, indicating that a significant fraction of the new spine synapses formed 1 h after LTP induction probably made mature and functional synapses.

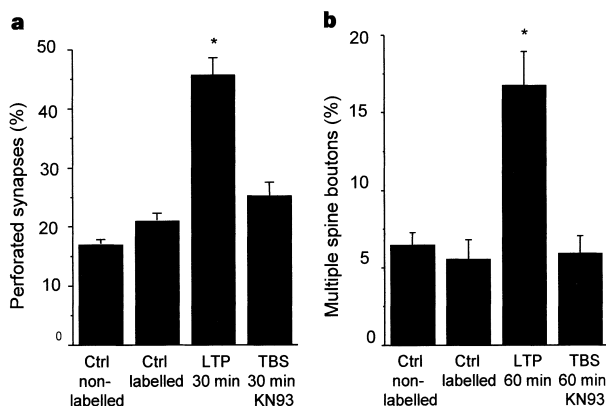


Figure 3 Blockade of LTP prevents morphological changes. **a**, Proportion of perforated spine profiles observed under control (ctrl) conditions (labelled and unlabelled spine profiles in non-stimulated cultures ($n = 3-4$, 319–630 spine profiles)), 30 min after LTP induction ($n = 4$, 358 spine profiles) and following high-frequency stimulation in the presence of $10 \mu\text{M}$ KN93 ($n = 4$, 313 spine profiles). **b**, Proportion of multiple spine boutons under control conditions (labelled and unlabelled spine profiles in non-stimulated cultures ($n = 3-4$, 319–630 spine profiles)), 60 min after LTP induction ($n = 4$, 308 spine profiles) and following high-frequency stimulation in the presence of $10 \mu\text{M}$ KN93 ($n = 3$, 183 spine profiles). The changes in both perforated synapses and multiple spine boutons were blocked by KN93 treatment (asterisk, $P < 0.001$ and $P < 0.01$, respectively).

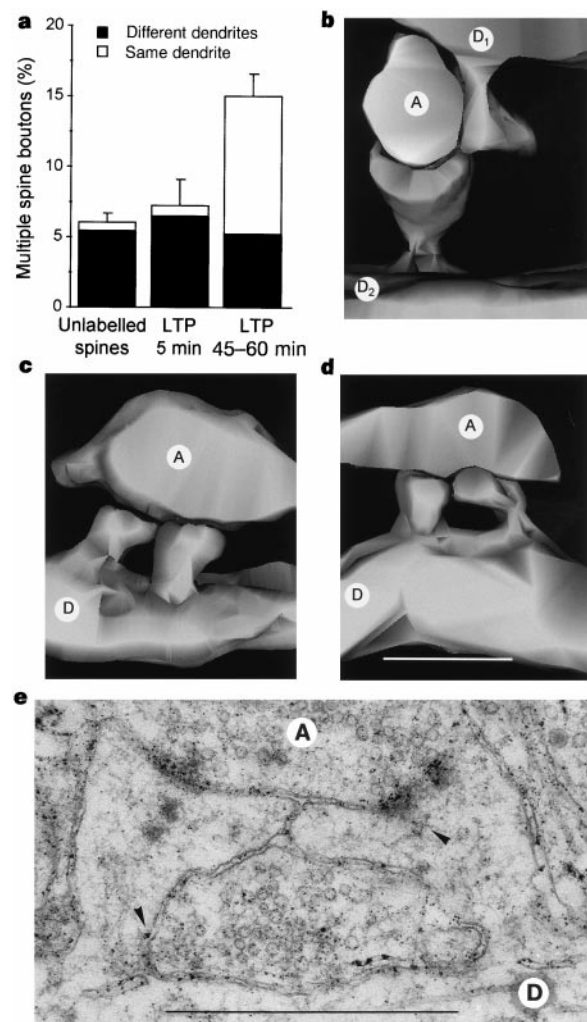


Figure 4 Three-dimensional reconstruction of multiple spine boutons. **a**, Proportion of multiple spine boutons measured on single sections. For each column, we calculated from three-dimensional reconstruction the proportion of boutons in which the two spines arose from different dendrites (black) or the same (white) dendrite (see text for numbers and proportions). **b**, A reconstructed, unlabelled multiple spine bouton in which the two spines arise from different dendrites. **c**, **d**, Labelled, multiple spine boutons observed 60 min after LTP induction in which the two spines arise from the same dendrite. Axons are indicated by A and dendrites by D. **e**, Two spines originating from the same dendrite (arrowheads: precipitate, unfortunately mainly present on another section). Scale bars, $1 \mu\text{m}$.

By using a calcium precipitation method to identify a subset of synapses likely to represent activated synapses, we have provided strong evidence that the late phase of LTP is associated with the formation of new, mature and functional spine synapses contacting the same presynaptic terminals.

One potential source of error in this study may come from selection biases. The calcium precipitate is primarily associated with smooth endoplasmic reticulum-like structures which are unevenly distributed among different types of synapse¹³. Furthermore, activity-dependent changes in the size of the spine head, PSD, perforation or smooth endoplasmic reticulum may affect quantitative results and particularly estimates of the proportion of perforated synapses^{1,2}. However, it seems unlikely that these possible biases would have modified our main conclusions. First, all data were obtained by analysing at different times the same population of synapses (labelled spine profiles), and strong evidence suggests that these mainly represented activated synapses⁴. Second, our main finding, the increase in the proportion of multiple spine boutons 60 min after LTP induction, was observed using single-section analyses, unbiased stereological methods and three-dimensional reconstruction. Third, these morphological changes were time-dependent and thus unrelated to the protocol used. Fourth, they were prevented by pharmacological blockade of LTP and thus are specifically associated with LTP. Finally, our conclusions are consistent with the results of many other studies^{5,6,14} and in particular with the demonstration by confocal microscopy of the occurrence of dendritic protrusions and spine-like structures following LTP induction^{7,8}.

An important finding added by the present study is that the new spines appearing after LTP induction include adjacent, mature synapses that contact the same presynaptic terminal. Note, however, that spine density was not directly assessed here and that new spines that do not share the same presynaptic bouton would not have been detected. These data nevertheless indicate that LTP may include duplication of synaptic contacts between an activated axon and a dendrite. Furthermore, the analysis of the presence of calcium precipitates in the spines arising from the same dendrite indicates that at least some of these newly formed synapses were functional. This mechanism could thus contribute to the stability of LTP and at the same time account for its synapse-specificity. From a quantitative point of view, our data indicate that newly formed synapses could represent up to 20% of stimulated synapses. This value is likely to be an underestimate, as newly formed synapses not sharing the same presynaptic bouton could not be detected.

Could new synapses have also been formed with other non-stimulated axon terminals? The increase in multiple spine boutons observed following LTP induction only concerned labelled spines and not unlabelled spines. Furthermore, we did not detect any increase in multiple spine boutons in which the two spines arose from different dendrites (the only situation which could eventually have reflected nonspecific synaptogenesis). Other studies, however, will be required to clarify this point.

The time course of the morphological changes is also of interest. The multiple spine boutons reported here were first detected 45–60 min after induction, a time course which is consistent with the occurrence of the new spines observed by Engert and Bonhoeffer⁷. Maletic-Savatic *et al.*⁸, however, reported a rapid occurrence of dendritic protrusions (within a few minutes) following high-frequency stimulation. We did not find such changes, although only structures defined by a clearly identified postsynaptic density were analysed. They could have thus remained undetected if the dendritic protrusions observed rapidly after stimulation did not exhibit all features of mature synapses^{15,16}. The most rapid change observed here was a transient increase in perforated synapses. How this morphological change relates to the dendritic protrusions and the subsequent formation of multiple spine boutons remains unclear^{10,17}. Perforated synapses could represent an intermediate step in the process of synapse formation and spine splitting^{18,19}. They

could also simply reflect the membrane dynamics and receptor turnover occurring at that time at activated synapses^{20–22}. It is of interest that the sequence of events reported here at the ultrastructural level coincides with the changes reported to occur with synapse turnover and formation of new synaptic contacts in several other *in vitro* models^{18,19,23,24}.

Together, these results indicate that the late phase of LTP is likely to be associated with the formation of new and possibly functional synapses between an activated terminal and its target cell. This mechanism of synapse duplication maintains specificity in information processing and provides a framework for understanding the numerous *in vivo* experiments that have demonstrated an increase in synapse number and complexity of dendritic arborization following exposure to an enriched environment or learning paradigms^{25,26}. □

Methods

Culture preparation and stimulation

Experiments were performed on hippocampal organotypic cultures prepared as described²⁷ and maintained for 12 days in culture. LTP was induced using TBS, a pattern of five bursts of stimulation at 5 Hz, each burst composed of four pulses at 100 Hz, repeated twice at a 10-s interval. To relabel activated synapses at the end of the experiments, a second theta burst was re-applied to the same group of afferents 5 min before fixation¹.

Morphological analyses

Slices were fixed for electron microscopy as described¹. Serial sections (ultratome Ultracut-E, Reichert-Jung) were stained *en bloc* first with 0.5% uranyl acetate in H₂O (pH 3.9) for 20 min before dehydration, then with 30 min of 2% uranyl acetate in ethanol after dehydration. Five to six sections per culture were usually examined with a Philips CM10 electron microscope at 80 kV. Synapses were randomly photographed at a magnification of 28,500× in an area corresponding to the middle portion of the apical arborization of CA1 pyramidal neurons. Images were digitized and dendritic spine profiles were classified blind by at least two independent observers as either labelled or unlabelled using as criteria the presence of accumulated electron-dense calcium precipitates in the spine head profile. In addition, we determined by densitometry the area of precipitate per spine profile (precipitate density; Fig. 1c). The results of the analyses made by the two observers and by densitometry agreed, with the exception of a small number of profiles (1–2%) that were discarded from quantitative analyses (white column, Fig. 1c). Synapses were identified by the presence of a definite postsynaptic membrane thickening and at least three presynaptic vesicles. Perforated synapses were identified by the presence of a discontinuity in the PSD and multiple spine boutons by the presence of at least two separate dendritic spines contacting the same presynaptic bouton. For three-dimensional analyses, ribbons of up to 60 serial sections were cut from each culture. Synaptic profiles corresponding to multiple spine boutons in which spines were clearly labelled or unlabelled were identified on the test section and then photographed serially at a magnification of at least 16,000×. For each slice culture, 3–13 labelled or unlabelled multiple spine boutons were reconstructed and the proportion of boutons in which the two spines arose from the same dendrite was calculated. For the stereological part of the study, the disector procedure was carried out as described²⁸. The proportion of labelled perforated synapses and multiple spine boutons was determined by systematic random sampling on six disector pairs in each block, using calcium precipitate as the counting unit.

All results are presented as mean $\pm$ s.e.m., with *n* indicating the number of slice cultures analysed. Statistical analyses were carried out using Student's *t*-test or the Mann–Whitney test.

Received 10 September; accepted 6 October 1999.

- Calverley, R. K. & Jones, D. G. Contributions of dendritic spines and perforated synapses to synaptic plasticity. *Brain Res. Brain Res. Rev.* **15**, 215–249 (1990).
- Geinisman, Y., Detolledo-Morrell, L. & Morrell, F. Induction of long-term potentiation is associated with an increase in the number of axospinous synapses with segmented postsynaptic densities. *Brain Res.* **566**, 77–88 (1991).
- Geinisman, Y. *et al.* Structural synaptic correlate of long-term potentiation: formation of axospinous synapses with multiple, completely partitioned transmission zones. *Hippocampus* **3**, 435–445 (1993).
- Buchs, P. A. & Muller, D. Induction of long-term potentiation is associated with major ultrastructural changes of activated synapses. *Proc. Natl. Acad. Sci. USA* **93**, 8040–8045 (1996).
- Geinisman, Y., Detolledo-Morrell, L., Morrell, F., Persina, I. S. & Beatty, M. A. Synapse restructuring associated with the maintenance phase of hippocampal long-term potentiation. *J. Comp. Neurol.* **368**, 413–423 (1996).
- Bolshakov, V. Y., Golan, H., Kandel, E. R. & Siegelbaum, S. A. Recruitment of new sites of synaptic transmission during the cAMP-dependent late phase of LTP at CA3-CA1 synapses in the hippocampus. *Neuron* **19**, 635–651 (1997).
- Maletic-Savatic, M., Malinow, R. & Svoboda, K. Rapid dendritic morphogenesis in CA1 hippocampal dendrites induced by synaptic activity. *Science* **283**, 1923–1927 (1999).
- Engert, F. & Bonhoeffer, T. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature* **399**, 66–70 (1999).
- Bliss, T. V. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
- Sorra, K. E., Fiala, J. C. & Harris, K. M. Critical assessment of the involvement of perforations, spinules, and spine branching in hippocampal synapse formation. *J. Comp. Neurol.* **398**, 225–240 (1998).

11. Buchs, P. A., Stoppini, L., Parducz, A., Siklos, L. & Muller, D. A new cytochemical method for the ultrastructural localization of calcium in the central nervous system. *J. Neurosci. Methods* **54**, 83–93 (1994).
12. Sorra, K. E. & Harris, K. M. Occurrence and three-dimensional structure of multiple synapses between individual radiatum axons and their target pyramidal cells in hippocampal area CA1. *J. Neurosci.* **13**, 3736–3748 (1993).
13. Spacek, J. & Harris, K. M. Three-dimensional organization of smooth endoplasmic reticulum in hippocampal CA1 dendrites and dendritic spines of the immature and mature rat. *J. Neurosci.* **17**, 190–203 (1997).
14. McKinney, R. A., Capogna, M., Durr, R., Gahwiler, B. H. & Thompson, S. M. Miniature synaptic events maintain dendritic spines via AMPA receptor activation. *Nature Neurosci.* **2**, 44–49 (1999).
15. Fiala, J. C., Feinberg, M., Popov, V. & Harris, K. M. Synaptogenesis via dendritic filopodia in developing hippocampal area CA1. *J. Neurosci.* **18**, 8900–8911 (1998).
16. Ziv, N. E. & Smith, S. J. Evidence for a role of dendritic filopodia in synaptogenesis and spine formation. *Neuron* **17**, 91–102 (1996).
17. Trommald, M. & Hulleberg, G. Dimensions and density of dendritic spines from rat dentate granule cells based on reconstructions from serial electron micrographs. *J. Comp. Neurol.* **377**, 15–28 (1997).
18. Carlin, R. K. & Siekevitz, P. Plasticity in the central nervous system: do synapses divide? *Proc. Natl Acad. Sci. USA* **80**, 3517–3521 (1983).
19. Nieto-Sampedro, M., Hoff, S. F. & Cotman, C. W. Perforated postsynaptic densities: probable intermediates in synapse turnover. *Proc. Natl Acad. Sci. USA* **79**, 5718–5722 (1982).
20. Fischer, M., Kaech, S., Knutti, D. & Matus, A. Rapid actin-based plasticity in dendritic spines. *Neuron* **20**, 847–854 (1998).
21. Shi, S. H. *et al.* Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* **284**, 1811–1816 (1999).
22. Liao, D., Zhang, X., O'Brien, R., Ehlers, M. D. & Huganir, R. L. Regulation of morphological postsynaptic silent synapses in developing hippocampal neurons. *Nature Neurosci.* **2**, 37–43 (1999).
23. Woolley, C. S., Wenzel, H. J. & Schwartzkroin, P. A. Estradiol increases the frequency of multiple synapse boutons in the hippocampal CA1 region of the adult female rat. *J. Comp. Neurol.* **373**, 108–117 (1996).
24. Kirov, S. A., Sorra, K. E. & Harris, K. M. Slices have more synapses than perfusion-fixed hippocampus from both young and mature rats. *J. Neurosci.* **19**, 2876–2886 (1999).
25. Moser, M. B., Trommald, M. & Andersen, P. An increase in dendritic spine density on hippocampal CA1 pyramidal cells following spatial learning in adult rats suggests the formation of new synapses. *Proc. Natl Acad. Sci. USA* **91**, 12673–12675 (1994).
26. Kleim, J. A., Vij, K., Ballard, D. H. & Greenough, W. T. Learning-dependent synaptic modifications in the cerebellar cortex of the adult rat persists for at least four weeks. *J. Neurosci.* **17**, 717–721 (1997).
27. Stoppini, L., Buchs, P. A. & Muller, D. A simple method for organotypic cultures of nervous tissue. *J. Neurosci. Methods* **37**, 173–182 (1991).
28. Sterio, D. C. The unbiased estimation of number and sizes of arbitrary particles using the disector. *J. Microsc.* **134**, 127–136 (1984).

Acknowledgements

We thank L. M. Cruz-Orive for advice on stereology; K. Harris for 3D reconstruction software; D. Smithies for morphometry software on AVS; L. Parisi and M. Moosmayer for culture preparation and technical assistance; and F. Pillonel for photographic work. This work was supported by the Swiss National Science Foundation, the Human Frontier Science Program, the National Priority Program and the Jean-Falk Vairant Foundation.

Correspondence and requests for materials should be addressed to D.M. (e-mail: Dominique.Muller@medecine.unige.ch).

Conserved regulation of proximodistal limb axis development by *Meis1/Hth*

Nadia Mercader*, Esther Leonardo*, Natalia Azpiazu†, Antonio Serrano*, Ginés Morata†, Carlos Martínez-A* & Miguel Torres*

* Departamento de Inmunología y Oncología, Centro Nacional de Biotecnología, CSIC-UAM, E-28049 Madrid, Spain

† Centro de Biología Molecular, CSIC-UAM, Universidad Autónoma de Madrid, E-28049, Madrid, Spain

Vertebrate limbs grow out from the flanks of embryos, with their main axis extending proximodistally from the trunk. Distinct limb domains, each with specific traits, are generated in a proximal-to-distal sequence during development¹. Diffusible factors expressed from signalling centres promote the outgrowth of limbs and specify their dorsoventral and anteroposterior axes^{2–4}. However, the molecular mechanism by which limb cells acquire

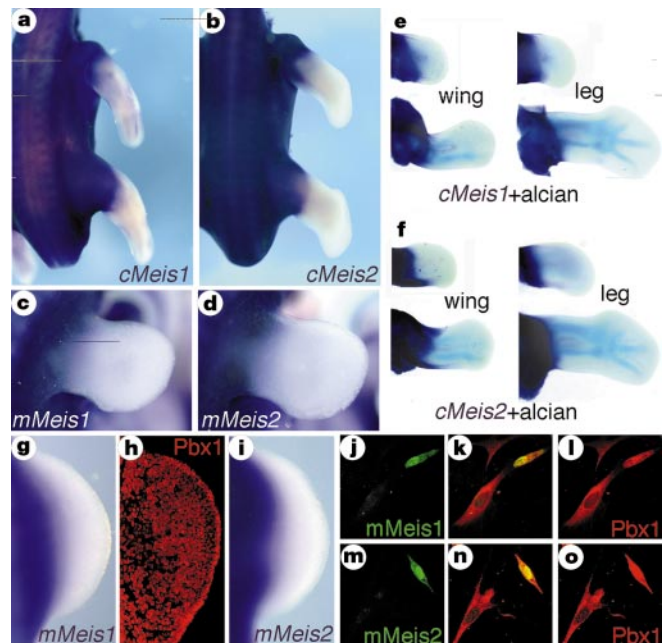


Figure 1 Coordinated *Meis/Pbx* expression in the limb bud. Whole-mount *in situ* hybridization (ISH) shows the expression of *cMeis1* (a), *cMeis2* (b), *mMeis1* (c) and *mMeis2* (d) in St. 29 chicken and embryonic day (E) 11.5 mouse limbs, respectively. Double alcian blue and ISH show the limits of *cMeis1* (e) and *cMeis2* (f) expression, together with chondrogenic condensations, between St. 23 and 28 of wing and leg development. **g–i**, Comparison of the expression domains of *mMeis1* (g) and *mMeis2* (i) messenger RNA with the nuclear domain of *Pbx1* (h). Endogenous *Pbx1* localization in mouse NIH 3T3 fibroblasts is shown in red (l, o) and HA-epitope detection denoting the presence and subcellular localization of *Meis1* and *Meis2* is shown in green (j, m). Co-localization of both proteins in the nucleus is shown in yellow (k, n).

their proximodistal (P–D) identity is unknown¹. Here we describe the role of the homeobox genes *Meis1/2* and *Pbx1* in the development of mouse, chicken and *Drosophila* limbs. We find that *Meis1/2* expression is restricted to a proximal domain, coincident with the previously reported domain in which *Pbx1* is localized to the nucleus⁵, and resembling the distribution of the *Drosophila* homologues *homothorax* (*hth*)^{5,6} and *extradenticle* (*exd*)⁷; that *Meis1* regulates *Pbx1* activity by promoting nuclear import of the *Pbx1* protein; and that ectopic expression of *Meis1* in chicken and *hth* in *Drosophila* disrupts distal limb development and induces distal-to-proximal transformations. We suggest that restriction of *Meis1/Hth* to proximal regions of the vertebrate and insect limb is essential to specify cell fates and differentiation patterns along the P–D axis of the limb.

Three distinct regions develop in a proximal-to-distal sequence along the vertebrate main limb axis; the stylopod (upper extremity), the zeugopod (lower extremity) and the autopod (hand or foot). The apical ectodermal ridge (AER), a specialized structure at the distal limb bud, and the zone of polarizing activity (ZPA), immediately posterior to the AER, produce signals that induce proliferation of progress zone cells, thereby promoting the growth of the limb bud and the generation of its main axis. Diffusible factors of the fibroblast growth factor (FGF) family emanating from the AER, and Sonic hedgehog from the ZPA, mediate the roles of these structures in P–D limb growth^{2,3}. Progress zone cells are stimulated by AER signalling to proliferate and generate the P–D limb axis by sequentially contributing progressively more distalized cells⁸.

We have studied the role of homeobox-containing genes showing conserved expression in specific P–D limb domains. Two highly related genes, *mMeis1* (refs 9, 10) and *mMeis2* (refs 11, 12), and the *Drosophila* orthologue *hth*⁶, are expressed in proximal domains of the limb primordium (Fig. 1c, d). To study the conservation of *Meis*

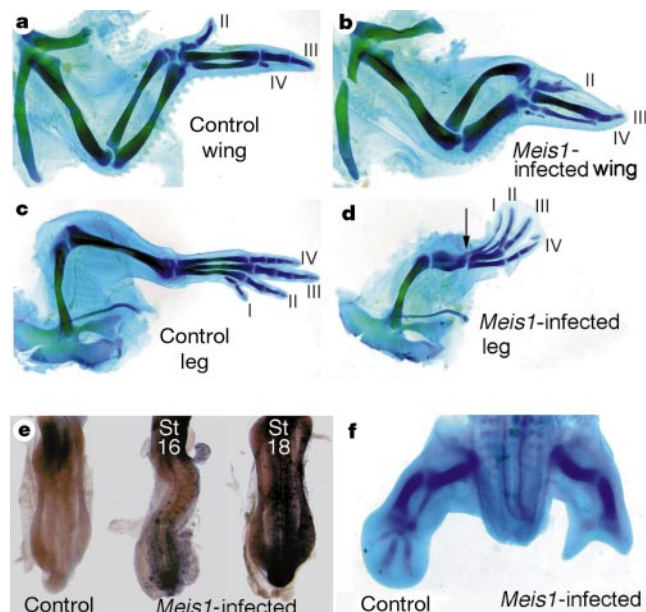


Figure 2 *Meis1* overexpression disrupts distal limb development in the chick. Skeletal preparations show control limbs (a, c) and limbs that were *Meis1*-infected between St. 10 and 13 (b, d). Arrow in d shows the tibial extension. e, *In situ* hybridization shows the extent of *RCAS-Meis1* expression in the limb anlage and early limb bud after infection at St. 4–8. f, Skeletal preparation shows an example of the strongest phenotype observed after early *RCAS-Meis1* infection.

genes among vertebrates, we isolated full-length coding regions for the corresponding chick orthologues, *cMeis1* and *cMeis2*, which are 98% and 99% conserved, with respect to their mouse counterparts. Expression of both genes in both species was detected in the trunk lateral mesoderm and extended continuously into a proximal limb bud domain, ending at the stylopod–zeugopod boundary (Fig. 1a–f).

In flies, *hth* function is coupled to that of the related homeobox protein Exd. *hth* controls *exd* function by promoting transport of its product to the nucleus^{13–16}. Mouse *Pbx* (refs 17, 18) activity is also regulated at the subcellular level. It is expressed in a proximal region of mouse limbs⁵ coincident with the *mMeis1/2* expression domain (Fig. 1g–i). To verify whether the subcellular regulation of Exd by Hth found in *Drosophila* is conserved in vertebrates, we overexpressed *mMeis1* and *mMeis2* genes in mouse fibroblasts, in which endogenous *Pbx1* localizes in the cytoplasm (Fig. 1l, o). Transfected cells containing nuclear *mMeis1/2* expression invariably had *Pbx1* in their nuclei (Fig. 1j–o). This indicates that *mMeis1/2* promotes nuclear localization of *Pbx1* and argues for conservation of the control mechanism.

We have investigated *Meis* function in vertebrate limbs by retroviral overexpression of murine *Meis1* in chicken embryos. Embryos infected with *RCAS-Meis1* showed altered limb development at a high frequency ($n = 158/272$ (positive/analysed)). Limbs infected between Hamburger–Hamilton stages (St.) 10 and 13 were greatly reduced in size, a phenotype evident only after St. 28. Skeletal preparations of St. 30–35 infected limbs showed that size reduction always involved the zeugopod and autopod ($n = 57/75$). Only very rarely ($n = 4$) was the stylopod slightly reduced (Fig. 2a–d). Given that this could be due to the earlier development of this region as compared with more distal regions, we performed *RCAS-Meis1* infections between St. 4 and 8 to achieve overexpression of *Meis1* between St. 16 and 18, during which the stylopod region is determined^{18,19}. Despite extensive *Meis1* expression in the limb anlage at St. 16, and in the limb bud at St. 18 (Fig. 2e), early infection did not affect stylopod development, whereas the autopod and zeugopod still were strongly affected ($n = 8$). Even

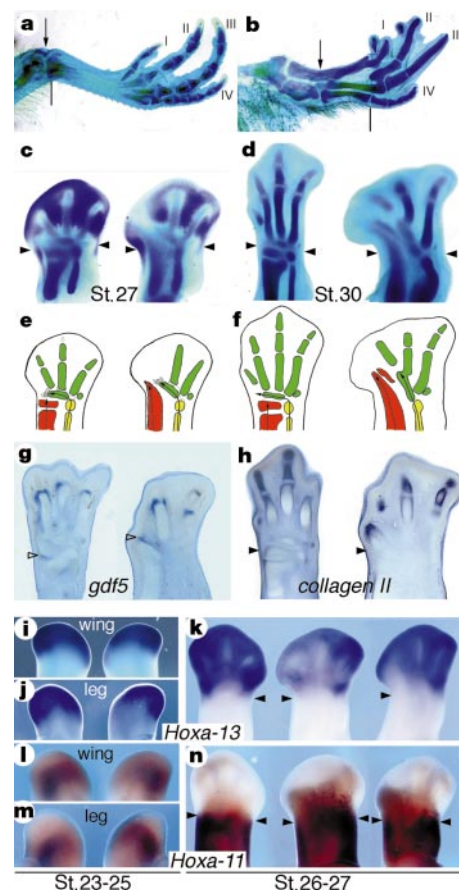


Figure 3 Evidence for the proximalization of skeletal elements in *Meis1*-infected limbs. Skeletal preparation of St. 38 (a, b), St. 27 (c) and St. 30 (d) *Meis1*-infected limbs, showing changes in the pattern of chondrogenesis. Roman numerals identify digits. In all pictures, infected limbs are shown on the right and control limbs on the left. Limbs in c were double-stained for alcian blue and *msx1*. e, f, Interpretative diagrams of the skeletal alterations shown in c and d, respectively. Tibia and tibial condensations are shown in red, fibula and fibulare in yellow and digital arch in green. Expression of *gdf5* (g) and type II collagen (h) is shown in longitudinal sections of control and infected St. 30 limbs; empty arrowheads in g indicate the expression of *gdf5* at the distalmost end of the tibia, to show its displacement in infected limbs. *Hoxa-13* expression is shown in early wings (i) and legs (j) and St. 26–27 legs (k). *Hoxa-11* expression is shown in early wings (l) and legs (m) and St. 26–27 legs (n). k, n, Single control limb on the left and two examples of infected limbs on the right. Filled arrowheads throughout, and arrows in a and b, indicate the anatomical limit between zeugopod and autopod, which corresponds in k and n to the limit between *Hoxa-13* and *Hoxa-11* expression in control limbs. Vertical bars (a, b) show the displacement of the distal limit of feather formation.

in the strongest case, in which early infection led to profound changes in zeugopod development and complete disruption of the autopod, the stylopod was unaffected (Fig. 2f).

The skeletal patterning of infected limbs was modified in several ways (see below), most strikingly in the anterior region of the leg. There, the tibia extended continuously through the ankle region into the autopod, spanning the region in which digit I and tarso-metatarsal II normally develop ($n = 22/55$) (Fig. 3a, b). To determine the origin of this alteration, we performed detailed analysis of early stages of chondrogenesis. Up to St. 28, *Meis1*-infected limbs showed no obvious alterations in overall limb size (Fig. 3c, k, n); however, changes in the chondrogenic condensations were already evident. During normal chondrogenesis, the tibial condensation stops at the autopod–zeugopod boundary and the digital arch develops in a posterior-to-anterior sequence from the fibular condensation. In infected limbs, the tibial chondrogenic centre extended ectopically into the autopod, invading the area in which

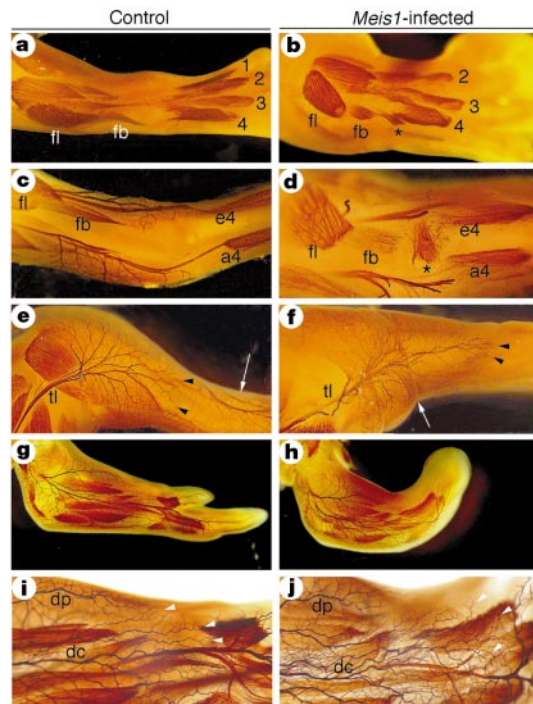


Figure 4 Alterations in distal nerve and muscle patterning in *Meis1*-infected limbs. Muscle and nerve patterns were identified by whole-mount immunohistochemistry in St. 34 limbs. Except in **a** and **b**, in which only MF20 was used, double staining was applied to reveal muscle and nerve patterns simultaneously. **a, b**, Dorsal leg views show alterations in autopod intrinsic digit extensors (labelled 1–4) and the extra muscle developing at the autopod dorsoposterior region (asterisk). **c, d**, Lateral leg views show a detail of the extra muscle (asterisk), placed between the extensor (e4) and abductor (a4) of digit 4 and showing the same fibre orientation as the fibularis longus (fl) (skin was removed for better observation of the fibre pattern). fb: fibularis brevis. **e, f**, Medial leg views show the tibialis lateralis nerve (tl) and its terminations (black arrowheads). White arrows indicate autopodal–zeugopodal boundaries. **g–h**, Dorsal wing views and enlargements of their autopodal–zeugopodal boundaries (**i, j**) show the increased muscular mass in the autopod region and its innervation by branches (white arrowheads) of the zeugopodal nerves dorsal propatagial (dp) and dorsal cutaneous (dc).

the most anterior region of the digital arch normally develops (Fig. 3c–f). The digital arch is thus posterodistally displaced and the condensation of its most anterior elements is inhibited or delayed (Fig. 3c–f). Analysis of the joint-formation marker *gdf5* (Fig. 3g) and the chondrogenesis marker type-II collagen (Fig. 3h) show no general disruption of their expression, but do show changes in their distribution that correlate with the altered disposition of skeletal elements in *Meis1*-infected limbs.

These pattern modifications indicated that ectopic *Meis1* expression might alter the normal sequence of generation of P–D identities in the limb. To test this, we examined the expression of the zeugopod-specific marker *Hoxa-11* and of the autopod-specific marker *Hoxa-13* (ref. 20). Infected limbs showed a distal shift of the *Hoxa-11* expression domain, which extended ectopically into the autopod (Fig. 3n) ($n = 16/26$). In parallel, the *Hoxa-13* proximal expression limit shifted distally and/or its expression level was generally downregulated in the autopod (Fig. 3k) ($n = 12/21$). In most specimens, the prospective digit I and tarso-metatarsal II areas showed ectopic *Hoxa-11* expression and *Hoxa-13* downregulation, coinciding with the region in which the tibial extension is found (Fig. 3k, n). The alterations observed in the expression of *Hoxa-11* and *Hoxa-13* first appeared around St. 23–25, when cells that later contribute to the boundary between autopod and zeugopod are being generated and acquire their identity in the progress zone^{8,19} (Fig. 3i, j, l, m). The normal sequence of distalization of progress

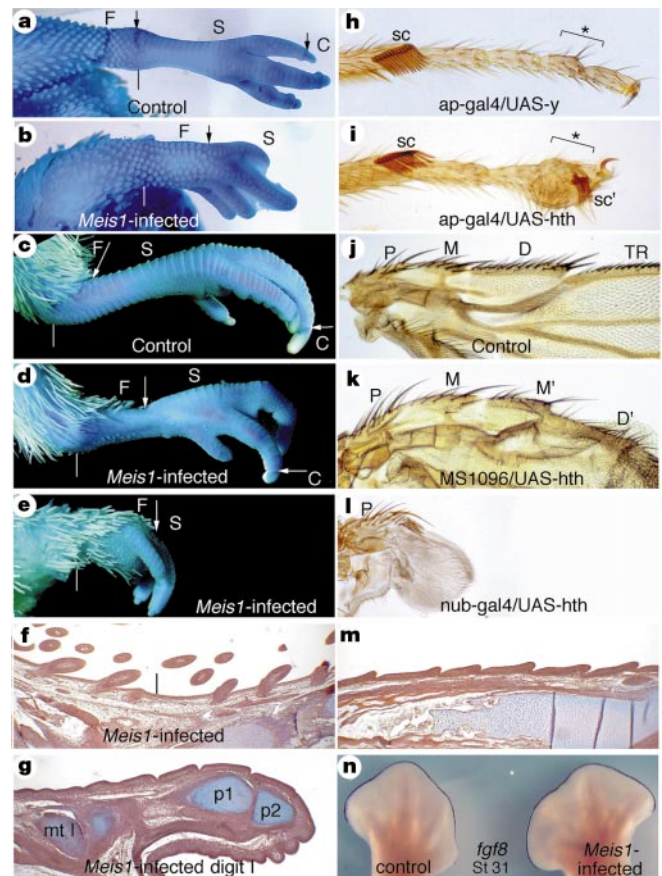


Figure 5 Ectodermal proximalization. **a–e**, *Meis1*-infected chicken limbs; **h–l**, ectopically *hth*-expressing *Drosophila* appendages. St. 37 (**a, b**) and St. 39 (**c–e**) alcian-blue-stained legs are shown. Throughout the Figure, arrows show the limits between the areas covered by the different skin structures, feathers (F) and scales (S) and claws (C), and straight line bars (black or white) indicate the anatomical limit between autopod and zeugopod, which was determined by bone inspection after clearing the embryos in methyl salicylate. Distal is always to the right except in **n**, in which distal is up. **f, g, m**, Histological sections of infected legs showing feathers extending through the limit between zeugopod and autopod (**f**), scales restricted to the autopodal region (**m**), and the lack of claw at the tip of digit I (**g**). mt I: metatarsal I; p1: phalange I; p2: phalange II. **n**, *In situ* hybridization showing *fgf8* expression. **h**, Distal region of the foreleg of a *Drosophila* male of the *ap-gal4/UAS-y* genotype. The pattern is normal, but the fourth tarsal segment is darker (asterisk), indicating the zone of *apterous* expression. Note the sex comb teeth (SC), characteristic of the basitarsus. **i**, Distal region of an *ap-gal4/UAS-hth* male foreleg. Note the additional sex comb teeth in the fourth tarsal segment, and the abnormal polarity of bristles in the zone extending beyond the area of *hth* expression. **j**, Anterior part of a *Drosophila* wing indicating the proximal (P), medial (M) and distal (D) Costa region and part of the triple row (TR) region. **k**, Anterior part of a wing of a *MS1096/UAS-hth* fly showing transformations of the distal Costa to medial Costa (M') and triple row to distal Costa (D'). **l**, Wing of a *nub-gal4/UAS-hth* fly showing that most of the appendage is lacking; only the most proximal elements remain.

zone cells thus appears to be altered in *Meis1*-expressing limbs, and the extension of the tibial condensation may result from recruitment of autopodal areas, mis-specified to zeugopodal identity.

In addition to the early effects on P–D patterning described, *Meis1* overexpression disrupts the patterning and growth of limb distal parts during the differentiation phase that takes place following the progress zone-dependent limb bud growth. Bone differentiation was delayed throughout the autopod and zeugopod, and patterning was generally disrupted in the autopod, which displayed fusion between consecutive phalanges and neighbouring digits, as well as interdigital extra-ossifications (Figs 2b, d, 3b). In addition, infected limbs showed persistence of interdigital membranes, which

correlated with reduced *msx1* expression in the prospective interdigital region (Figs 2d, 3b, c). The timing of appearance and type of alterations seen are similar to those reported for several Hox gene gain-of-function and loss-of-function mutations^{21,22}. Given the alterations observed in *Hoxa* gene expression, at least some of the defects observed during the differentiation phase may result from the presence of a modified combination of Hox gene activity in differentiating distal limb regions.

Further evidence for proximalization of autopod structures was obtained by examining the muscle and nerve patterns of *Meis1*-infected legs. Four intrinsic extensor muscles develop in the dorsal tarso-metatarsal region of the leg autopod, each associated with its respective digit (Fig. 4a). In *Meis1*-infected legs, the digit I extensor, which should have developed in the area invaded by the enlarged tibia, was either absent or greatly reduced and displaced to the digit II area (Fig. 4b) ($n = 11/12$). In addition, this autopodal area was innervated by branches of the tibialis lateralis, a nerve that normally extends over the tibia region, never crossing the zeugopodal-autopodal boundary (Fig. 4e, f) ($n = 8/8$). Furthermore, infected limbs developed an extra muscle at the proximal end of tarso-metatarsus IV (Fig. 4b-d) ($n = 8/12$). The shape and fibre orientation of this autopodal extra muscle resembled those of the fibularis longus, a zeugopodal muscle situated at the proximal head of the fibula (Fig. 4d). Similar results were obtained for the wing autopod ($n = 7/11$), which was innervated by branches of the dorsal propatagial and dorsal cutaneous zeugopodal nerves (Fig. 4i, j) and showed an increased muscle mass (to which, however, identity could not be assigned; Fig. 4g, h).

Alterations in P-D patterning were also seen in skin elements ($n = 17/25$). Three different ectodermal derivatives develop along the P-D axis on chicken legs: feathers on the stylopod and zeugopod, scales on the autopod and claws at the digit tips (Fig. 5a, c). Infected limbs showed feathers extending continuously from the zeugopod onto the autopod, up to the digit insertion area (Fig. 5b, d-f, m; see also Fig. 3b). Scales are replaced by feathers in the proximal autopod and become restricted to the digits (Fig. 5b, e) and the two-row scale pattern typical of the tarso-metatarsal region extends distally into the toes, replacing their normal one-row scale pattern (Fig. 5b). The most distal structures, the claws, were either reduced or completely lost, and were replaced by scales (Fig. 5b, d, e; see also Fig. 3b). This last phenotype is not due to premature termination of AER activity, as it was observed in digits in which the last phalange is formed (Fig. 5g) ($n = 15/29$), and *Meis1*-expressing limbs do not show premature loss of AER function, as revealed by normal *fgf8* expression in late stages (Fig. 5n) ($n = 12$).

Our results show that the activity of *Meis1* prevents distal limb development by inhibiting growth and differentiation (a similar conclusion was reached for *Meis2* (ref. 23)), and alters the normal P-D sequence of generation of cell identities by inhibiting distalization.

In *Drosophila*, *exd* and *hth* are required for proximal but not distal leg development, whereas their ectopic expression results in the elimination of the distal leg^{5-7,24}. Using the *gal4/UAS* method²⁵, we examined the patterning effects of ectopic expression of *hth* in the development of wings and legs of *Drosophila*. We used several *gal4* lines that direct expression in the appendages, and induced clones of *hth*-expressing cells. The results obtained closely resemble those described above for the chicken limb. When *hth* is driven by *gal4* lines conferring expression in the distal regions of the appendages, for example, *Dll-gal4* in legs and *ap-* or *nub-gal4* in wings, appendage development is aborted (Fig. 5l; and data not shown). In cases in which ectopic expression is more restricted, such as *ap-gal4/UAS-hth* in leg, *MS1096/UAS-hth* in wing or in *hth*-expressing clones, *hth* induces changes in the fate of the distal cells, which differentiate into pattern elements corresponding to more proximal positions (Fig. 5h-k). In *ap-gal4/UAS-hth* males, the fourth tarsal segment of the first leg presents the characteristic sex comb teeth of

the basitarsus (Fig. 5h, i). In *MS1096/UAS-hth* flies, pattern elements of the wing margin, such as the triple row of bristles, differentiate like the more proximal Costa structures (Fig. 5j, k). Similarly, clones of *hth*-expressing cells in the leg differentiate into structures corresponding to more proximal positions, such as sex comb teeth in any of the tarsal segments or femur-like bristles in the tibia (not shown). Ectopic *hth* expression also induced changes in bristle differentiation and polarity that extended beyond the zone of *hth* expression (Fig. 5i), suggesting a non-autonomous effect on patterning; a similar observation has been reported²⁶. The conclusion that *hth* induces a shift towards more proximal fates is also supported by the observation²⁴ that *hth*-expressing clones in tarsal segments express *dachshund*, a marker of the femur and tibia, which are more proximal leg segments.

Meis1 in vertebrates is likely to act by inhibiting the proper sequence of distalization of progress zone cells in response to AER signals. Its confinement to proximal limb regions during normal development may thus protect cells in proximal compartments from being distalized by AER signals. Even though the mechanism of action may be different, the similarities observed indicate that some aspects of *Meis/Hth* and *Pbx1/Exd* function in P-D limb axis development are shared between insects and vertebrates. In both cases, expression of 'trunk genes' promotes proximal limb development, but must be repressed to allow distal development. This shared strategy may have appeared during evolution as a necessary step to allow outgrowth of ancient appendages from the trunk of a common ancestor, thereby generating a primordial molecular diversity along the P-D axis. During subsequent evolution, this molecular diversity may have been exploited to specify cell fate and differentiation programmes characteristic of different domains along the P-D limb axis. This shared mechanism supports the proposed view for a common origin of the regulatory pathways used during development of metazoan appendages as the basis for their molecular similarities²⁷. □

Methods

cMeis1 and *cMeis2* cloning

Oligonucleotides specific for *mMeis1* and *mMeis2* were used at low-stringency annealing temperatures to amplify *cMeis1* and *cMeis2* complete coding sequences from total chicken embryo complementary DNA.

Probes

Gene names and nucleotides used: *cmx1*: +312 to +1307; *cHoxa-11*: +616 to +1160; *cHoxa-13*: +487 to +1159; *mMeis1*: +1 to +1398; *mMeis2*: +1 to +1528; *cMeis1*: +439 to +1491; *cMeis2*: +532 to +999; *cgd5*: +805 to +1484; *cCollagen II*: +4 to +537; *cfgf8*: +1 to +645.

Antibodies

Rabbit polyclonal antibody Pbx1 (Santa Cruz Biotechnology), the mouse monoclonals anti-HA epitope and anti-NF68 (Sigma) and mouse monoclonal MF20 (University of Iowa hybridoma bank) were used as recommended by the suppliers.

Constructs and cell transfections

Coding regions of *mMeis1* and *mMeis2* were amino-terminally fused to the haemagglutinin (HA)-epitope and inserted in the polylinker of the PC-DNA3 expression vector (Invitrogen) under the control of the cytomegalovirus promoter. NIH 3T3 cells were lipofected as recommended by the manufacturer (Gibco-BRL). Retroviral infection²⁸, *in situ* hybridization²⁹, immunohistochemistry on mouse sections and cultured cells⁵ and skeletal preparations²² were done as described.

Drosophila work

Experiments to induce ectopic *hth* expression were based on the *gal4/UAS* method²⁵. *ap-gal4* is an insertion of the pGawB element in the *apterous* (*ap*) gene and drives expression in the dorsal wing compartment³⁰ and the fourth tarsal leg segment; *nub-gal4* is an insertion of the same transposon on the *nubbin* (*nub*) gene and directs expression in the wing pouch, and the *MS1096* line drives expression in most of the dorsal and ventral wing regions. *hth*-expressing cell clones were induced by heat shocking larvae of genotype *yf⁶⁶hsFLP*; *abx > f > gal4-lacZ/UAS-hth*; *UAS-y*. FLP-induced recombination produces clones of cells in the thoracic segments³¹ that express both *hth*⁺ and *y*⁺ and are also mutant for the cuticle marker *f⁶⁶*.

Received 21 September; accepted 20 October 1999.

- Johnson, R. L. & Tabin, C. J. Molecular models for vertebrate limb development. *Cell* **90**, 979–990 (1997).
- Niswander, L., Jeffrey, S., Martin, G. R. & Tickle, C. A positive feedback loop coordinates growth and patterning in the vertebrate limb. *Nature* **371**, 609–612 (1994).
- Laufer, E., Nelson, C. E., Johnson, R. L., Morgan, B. A. & Tabin, C. Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. *Cell* **79**, 993–1003 (1994).
- Parr, B. A. & McMahon, A. P. Dorsalizing signal Wnt-7a required for normal polarity of D–V and A–P axes of mouse limb. *Nature* **374**, 350–353 (1995).
- González-Crespo, S. *et al.* Antagonism between extradenticle function and Hedgehog signalling in the developing limb. *Nature* **394**, 196–200 (1998).
- Abu-Shaar, M. & Mann, R. S. Generation of multiple antagonistic domains along the proximodistal axis during *Drosophila* leg development. *Development* **125**, 3821–3830 (1998).
- González-Crespo, S. & Morata, G. Genetic evidence for the subdivision of the arthropod limb into cospote and telopodite. *Development* **122**, 3921–3928 (1996).
- Summerbell, D., Lewis, J. & Wolpert, L. Positional information in chick limb morphogenesis. *Nature* **224**, 492–496 (1973).
- Moskow, J. J., Bullrich, F., Huebner, K., Daar, I. O. & Buchberg, A. M. Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol. Cell. Biol.* **15**, 5434–5443 (1995).
- Nakamura, T., Jenkins, N. A. & Copeland, N. G. Identification of a new family of Pbx-related homeobox genes. *Oncogene* **13**, 2235–2242 (1996).
- Cecconi, F., Proetzel, G., Alvarez-Bolado, G., Jay, D. & Gruss, P. Expression of Meis2, a Knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. *Dev. Dyn.* **210**, 184–190 (1997).
- Oulad-Abdelghani, M. *et al.* Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. *Dev. Dyn.* **210**, 173–183 (1997).
- Berthelsen, J., Kilstrup-Nielsen, C., Blasi, F., Mavilio, F. & Zappavigna, V. The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes Dev.* **13**, 946–953 (1999).
- Abu-Shaar, M., Ryoo, H. D. & Mann, R. S. Control of the nuclear localization of extradenticle by competing nuclear import and export signals. *Genes Dev.* **13**, 935–945 (1999).
- Pai, C. Y. *et al.* The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in *Drosophila*. *Genes Dev.* **12**, 435–446 (1998).
- Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. & Mann, R. S. Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* **91**, 171–183 (1997).
- Kamps, M. P., Murre, C., Sun, X. & Baltimore, D. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell* **60**, 547–555 (1990).
- Nourse, J. *et al.* Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell* **60**, 535–545 (1990).
- Rowe, D. A. & Fallon, J. F. The proximodistal determination of skeletal parts in the developing chick leg. *J. Embryol. Exp. Morphol.* **68**, 1–7 (1982).
- Nelson, C. E. *et al.* Analysis of Hox gene expression in the chick limb bud. *Development* **122**, 1449–1466 (1996).
- Zakany, J. & Duboule, D. Hox genes in digit development and evolution. *Cell. Tissue Res.* **296**, 19–25 (1999).
- Morgan, B. A., Izpisua-Belmonte, J. C., Duboule, D. & Tabin, C. J. Targeted misexpression of Hox-4.6 in the avian limb bud causes apparent homeotic transformations. *Nature* **358**, 236–239 (1992).
- Capdevila, J., Tsukui, T., Rodríguez-Esteban, C., Zappavigna, V. & Izpisua-Belmonte, J. C. Regulatory interactions between the proximal determinant Meis2 and distal antagonism of BMP by Gremlin Control Vertebrate limb outgrowth. *Mol. Cell* (in press).
- Wu, J. & Cohen, S. M. Proximodistal axis formation in the *Drosophila* leg: subdivision into proximal and distal domains by Homothorax and Distal-less. *Development* **126**, 109–117 (1999).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Goto, S. & Hayashi, S. Proximal to distal cell communication in the *Drosophila* leg provides a basis for an intercalary mechanism of limb patterning. *Development* **126**, 3407–3413 (1999).
- Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
- Logan, M. & Tabin, C. Targeted gene misexpression in chick limb buds using avian replication-competent retroviruses. *Methods* **14**, 407–420 (1998).
- Wilkinson, D. G. & Nieto, M. A. Detection of messenger RNA by *in situ* hybridization to tissue sections and whole mount. *Methods Enzymol.* **225**, 361–373 (1993).
- Calleja, M., Moreno, E., Pelaz, S. & Morata, G. Visualization of gene expression in living adult *Drosophila*. *Science* **274**, 252–255 (1996).
- Celis, J. F. & Bray, S. Feed back mechanisms affecting Notch activation at the dorsoventral boundary in the *Drosophila* wing. *Development* **124**, 3241–3251 (1997).

Acknowledgements

We thank P. Bovolenta and Á. Nieto for advice and help during the development of this project; Á. Nieto, D. R. Jones and C. Mark for critical reading of the manuscript; J. C. Izpisua-Belmonte for unpublished results; and N. Copeland, J. C. Izpisua-Belmonte, R. S. Mann and C. Tabin for probes. N.M. is a recipient of an ETH-CSIC collaborative fellowship and an EU Marie Curie predoctoral fellowship. N.A. and G.M. are supported by grants from the Spanish Dirección General de Investigación Científica y Técnica and Human Frontiers Program. The Department of Immunology and Oncology was founded and is supported by the Spanish Research Council (CSIC) and Farmacia & Upjohn.

Correspondence and requests for materials should be addressed to M.T. (e-mail: mtorres@cnb.uam.es). The sequences have been deposited in the GenBank under accession codes AF202933 (*cMeis1*) and AF202934 (*cMeis2*).

CtBP/BARS induces fission of Golgi membranes by acylating lysophosphatidic acid

Roberto Weigert^{*†}, Maria Giuseppina Silletta^{*†}, Stefania Spanò^{*}, Gabriele Turacchio^{*}, Claudia Cericola^{*}, Antonino Colanzi^{*}, Silvia Senatore^{*}, Raffaella Mancini^{*}, Elena V. Polishchuk^{*}, Mario Salmona[‡], Francesco Facchiano^{*§}, Koert N. J. Burger^{||}, Alexander Mironov^{*}, Alberto Luini^{*} & Daniela Corda^{*}

^{*} Department of Cell Biology and Oncology, Istituto di Ricerche Farmacologiche “Mario Negri”, Consorzio Mario Negri Sud, 66030 Santa Maria Imbaro (Chieti), Italy

[‡] Department of Biochemistry and Molecular Pharmacology, Istituto di Ricerche Farmacologiche “Mario Negri”, 20157 Milano, Italy
[§] Laboratory of Vascular Pathology, Istituto Dermatologico dell’Immacolata, Rome, Italy

^{||} Department of Biochemistry of Membranes, Institute of Biomembranes, Utrecht University, 3584 CH Utrecht, The Netherlands

[†] These authors contributed equally to this paper

Membrane fission is essential in intracellular transport. Acyl-coenzyme As (acyl-CoAs) are important in lipid remodelling and are required for fission of COPI-coated vesicles¹. Here we show that CtBP/BARS, a protein that functions in the dynamics of Golgi tubules², is an essential component of the fission machinery operating at Golgi tubular networks, including Golgi compartments involved in protein transport and sorting. CtBP/BARS-induced fission was preceded by the formation of constricted sites in Golgi tubules, whose extreme curvature is likely to involve local changes in the membrane lipid composition. We find that CtBP/BARS uses acyl-CoA to selectively catalyse the acylation of lysophosphatidic acid to phosphatidic acid both in pure lipidic systems and in Golgi membranes, and that this reaction is essential for fission. Our results indicate a key role for lipid metabolic pathways in membrane fission.

We first studied the mode of action of CtBP/BARS in *in vitro* assays of Golgi tubular network morphology using negative staining of whole-mount preparations and electron microscopy (EM). Golgi-enriched fractions appeared as groups of partially stacked disk-like cisternae fenestrated at the edges and covered with tubular-reticular networks, which often extended laterally from the cisternal margins (Fig. 1a). Their organization and size agree with those observed by the freeze-etch technique³, and with descriptions of the Golgi stacks in chemically fixed cells⁴. Moreover, they contain high levels of the Golgi protein giantin and of the *cis* Golgi proteins GM130 and p115 (ref. 5), as judged by immuno-gold labelling with specific antibodies (data not shown). The same membranes were also examined by resin embedding and thin sectioning for EM. They exhibited the expected stacked structures with two to four cisternae, accompanied by surrounding tubular domains (see Supplementary Information).

When these Golgi membranes were exposed to either recombinant or partially purified rat brain CtBP/BARS in 2 mg ml⁻¹ cytosol, the tubular structures shown in Fig. 1a underwent rapid disruption, resulting in the formation of vesicular fragments of irregular sizes (50–90 nm), which were sometimes aligned in rows, and in the appearance of naked (lacking tubular reticular networks) cisternae (Fig. 1b). Cytosol alone had similar, but much slower, effects. The extent and time course of the fission/fragmentation of Golgi tubular domains was quantified and is shown in Fig. 1c. There was a twofold increase in the number of vesicles and buds after 15 min in cytosol, whereas tubules were not affected. A clear breakdown of the tubular networks was apparent only after 30 min. When cytosol was immunodepleted of CtBP/BARS, the breakdown of tubular Golgi

domains was inhibited. The same inhibition was achieved both by pre-incubating the cytosol with an antibody against peptide 174–181 of CtBP/BARS (PEP9)², and by ADP-ribosylating CtBP/BARS (data not shown; CtBP/BARS is known to be ADP-ribosylated in the presence of brefeldin A, which results in loss of CtBP/BARS activity²). When depleted cytosol was supplemented with recombinant or partially purified rat brain CtBP/BARS, it regained its fission-inducing activity. At endogenous levels of partially purified CtBP/BARS (0.3 $\mu\text{g ml}^{-1}$), activity returned to control levels, and at a higher concentration (3 $\mu\text{g ml}^{-1}$) the fission/fragmentation process was greatly stimulated (Fig. 1c). Recombinant CtBP/BARS had identical effects, but was 5–10-fold less potent than the brain protein (Fig. 1d). Control incubations (2 mg ml⁻¹ soybean trypsin inhibitor) for 60 min did not alter the structure of Golgi membranes (Fig. 1c). These data show that CtBP/BARS is an essential rate-controlling component of the tubule fission-inducing activity of cytosol. However, CtBP/BARS alone (without cytosol) was inactive (Fig. 1c), indicating that it requires cytosolic cofactors to function.

The effects of CtBP/BARS on Golgi tubular networks were also examined using conventional resin embedding and thin sectioning; the results were in close agreement with the negative-stain data (see Supplementary Information). Finally, we examined whether, as a consequence of CtBP/BARS-induced fission of Golgi tubules, vesicular fragments are released from stacks. This release was studied both morphometrically (by counting the released fragments) and biochemically (by measuring the albumin and transferrin present in the fragments)⁶. In both these assays, CtBP/BARS strongly induced the formation and release of fragments, mirroring precisely the effects observed on the morphology of the Golgi stacks (see Supplementary Information), and confirming the fission-inducing action of CtBP/BARS.

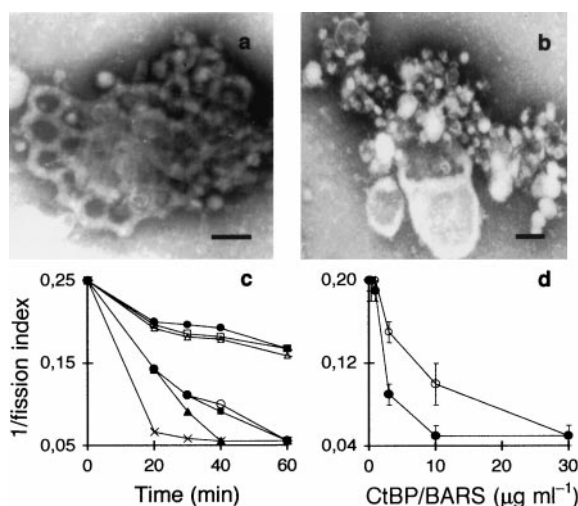


Figure 1 CtBP/BARS induces fission of the Golgi tubular network (negative staining). **a**, Untreated membranes. **b**, Golgi membranes incubated for 20 min with 2 mg ml⁻¹ cytosol supplemented with 3 $\mu\text{g ml}^{-1}$ partially purified CtBP/BARS. Identical effects occurred with 15 $\mu\text{g ml}^{-1}$ recombinant CtBP/BARS. Scale bar, 100 nm. **c**, Time course of CtBP/BARS-induced fission/fragmentation. Golgi membranes were incubated with 2 mg ml⁻¹ soybean trypsin inhibitor (open squares), 3 $\mu\text{g ml}^{-1}$ partially purified CtBP/BARS in the absence of cytosol (filled circles), 2 mg ml⁻¹ control cytosol (filled squares), 2 mg ml⁻¹ CtBP/BARS-depleted cytosol without additions (open triangles), or supplemented with 0.3 (open circles), 1.0 (filled triangles) or 3.0 (crosses) $\mu\text{g ml}^{-1}$ partially purified CtBP/BARS. See Methods for calculation of the fission index. Values are means of triplicates. The experiments were repeated 3–5 times with similar results; the s.d. never exceeded 15% of the mean. **d**, Fission/fragmentation activity of partially purified and recombinant CtBP/BARS. Golgi membranes were incubated for 20 min with 2 mg ml⁻¹ CtBP/BARS-depleted cytosol supplemented with increasing concentrations of either partially purified or recombinant CtBP/BARS. Values are means $\pm$ s.d. of three independent experiments.

A feature of Golgi membranes treated with CtBP/BARS-enriched cytosol was the early (within 10–15 min) appearance of a number of sites (9 $\pm$ 2 per stack) in which the tubule diameter was greatly reduced. Figure 2 shows a selection of such sites, visualized using both negative-stain and resin-embedded thin-sectioned samples. The degree of constriction was often extreme, with tubule diameters down to a value of 11 $\pm$ 1 nm. Constrictions were unevenly distributed among the tubules themselves, although when several constrictions were present in a single tubule, they were spaced at regular intervals (85 $\pm$ 3 nm). These dimensions are compatible with those of the vesicles produced by fission. Moreover, images were found of what appeared to be stages of transition between the series of constricted sites and the rows of aligned vesicles often observed at later times in the incubation. We propose that these are sites where fission occurs later, and will refer to them as ‘fission intermediates’ henceforth. Fission intermediates were also found in samples treated with cytosol alone, albeit rarely (2 $\pm$ 1 per stack). Thus, CtBP/BARS increases the frequency of a phenomenon that is also induced by normal cytosol. In resin-embedded, thin-sectioned samples, fission intermediates were of variable diameter, very similar to those seen by negative stain, with a minimum diameter of 12 $\pm$ 1 nm (from tubules of 48 $\pm$ 2 nm). The bilayer structure was visible and was the expected 4 nm, leaving only 4 nm for the diameter of the lumen.

At equilibrium, this level of membrane curvature is almost certainly incompatible with the lipid composition of a normal membrane bilayer. The minimum vesicular diameter that will accommodate ‘cylindrical’ lipids (for example, phosphatidylcholine) is 25 nm (ref. 7). Thus, the presence of cylindrical lipid species should be energetically unfavourable in constricted fission sites. Instead, an enrichment with ‘cone-shaped’ lipids (for example, diacylglycerol)⁷ in the internal leaflet, and ‘inverted cone-shaped’ lipids (lysolipids)⁷ in the external leaflet of the membrane should facilitate the formation of the highly curved intermediates, and reduce the energy barriers to fission. As CtBP/BARS requires cytosolic cofactors to induce fission, we reasoned that such factors might induce local changes in lipid metabolism. Long-chain acyl-CoAs are major acyl donors to lysolipids and lipid headgroups, and are abundant in cytosol⁸; moreover, they are required for fission of COPI-coated vesicles¹. To test whether they might be the cytosolic CtBP/BARS cofactors, CtBP/BARS was mixed with Golgi membranes without cytosol but in the presence of palmitoyl-CoA (pCoA). Unexpectedly, pCoA replaced cytosol in inducing the fission/fragmentation of Golgi tubules (Fig. 3a–c), and in releasing vesicular fragments from stacks (see Supplementary Information).

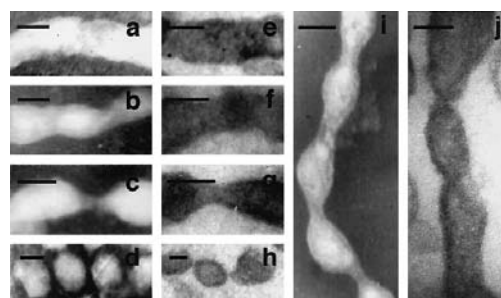


Figure 2 Structure of CtBP/BARS-induced fission intermediates in Golgi tubules. **a**, **b**, **d**, **i**, Negative stains of whole-mount preparations. **e**, **f**, **h**, **j**, Thin sections of resin-embedded samples. **a**, **e**, Membranes incubated for 15 min with control cytosol; examples of unconstricted tubules; **b**, **d**, **f**, **h**, Golgi membranes incubated for 15 min with 3 $\mu\text{g ml}^{-1}$ rat brain, or 30 $\mu\text{g ml}^{-1}$ recombinant CtBP/BARS added to 2 mg ml⁻¹ cytosol. Moderate constrictions are seen in **b** and **f**, and extreme constrictions in **c** and **g**. Rows of aligned vesicles, presumably resulting from clipping of a single tubule at regular intervals, are seen in **d** and **h**; series of fission intermediates in single tubules in **i** and **j**. Scale bar, 40 nm.

pCoA alone was without effect. Recombinant CtBP/BARS was 4–8-fold less potent than the partially purified protein (Fig. 3d), and ADP-ribosylated CtBP/BARS was nearly without effect (Fig. 3c). CtBP/BARS was slightly less potent (twofold) in causing tubule fragmentation, and constricted fission intermediates were more numerous (3–4-fold) than in the presence of cytosol. Perhaps the cytosol contains additional factors which destabilize fission intermediates to complete the fission process. The fission-inducing effects of CtBP/BARS were dependent on pCoA concentration and incubation time. Using $10 \mu\text{g ml}^{-1}$ recombinant CtBP/BARS and a time of 20–30 min, the threshold concentration and the maximal concentration of pCoA were $2 \mu\text{M}$ and $10 \mu\text{M}$, respectively. The effect of the length and degree of saturation of acyl chains in acyl-CoAs were then tested. The shortest active molecule was myristoyl-CoA (C14:0), and the potency increased with chain length and unsaturation. Arachidonoyl-CoA (C20:4) and oleoyl-CoA (C18:1) were the most potent (Fig. 4a, black bars). Notably, CoA itself was active at $50 \mu\text{M}$, but only in the presence of millimolar concentrations of ATP, whereas the acyl-CoAs did not require this nucleotide. This is probably because acyl-CoAs can be resynthesized from CoA-mediated free (non-esterified) fatty acids found in membranes, and the ligase catalysing this reaction requires millimolar ATP⁹.

These observations indicated that CtBP/BARS may be involved in acyl-CoA-dependent lipid modifications, and we tested this by incubating a series of lipids, lysolipids and their corresponding headgroups with recombinant CtBP/BARS and [¹⁴C]pCoA. Of these compounds, lysophosphatidic acid (LPA) was the only one that showed activity as an acyl acceptor, resulting in incorporation of label into a new species. This product was [¹⁴C]phosphatidic acid (PA), as demonstrated by co-elution with pure PA standard in thin-layer chromatography (TLC) (Fig. 4b) and high-performance liquid chromatography systems (data not shown). Thus, CtBP/BARS is an LPA-acyltransferase. In kinetic experiments, in $5 \mu\text{M}$ pCoA and

$30 \mu\text{M}$ LPA, the rate of PA formation was constant for 40 min and was proportional to the concentration of recombinant CtBP/BARS (Fig. 4c). These concentrations of LPA and pCoA were optimal, as beyond them the formation of PA became inhibited (see bell-shaped curves in Fig. 4d, e). This precludes the determination of the reaction K_m and K_{cat} . As most acyltransferases are membrane enzymes⁸, LPA was incorporated in liposomes prepared from a total lipid extract of purified Golgi fractions. Here, the PA yield increased fourfold (Fig. 4g), indicating that CtBP/BARS is more efficient when it acts in a membrane context; however, the substrate concentration–velocity curves remained bell-shaped (data not

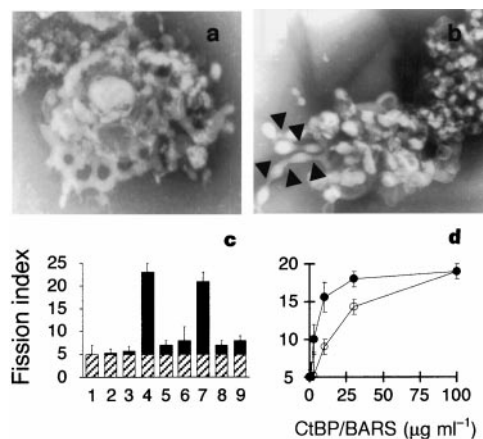


Figure 3 Acyl-CoAs are essential cofactors for CtBP/BARS-induced fission/fragmentation of Golgi tubular networks. **a, b**, Golgi membranes visualized by negative staining. Membranes were incubated for 20 min with $10 \mu\text{g ml}^{-1}$ partially purified CtBP/BARS alone (**a**) or in $10 \mu\text{M}$ pCoA (**b**). **c**, Quantification of the CtBP/BARS fission/fragmentation effects. Untreated membranes (1). Golgi membranes incubated for 15 min with no additions (2), cytosol (3), cytosol supplemented with $10 \mu\text{g ml}^{-1}$ partially purified CtBP/BARS (4), $10 \mu\text{g ml}^{-1}$ partially purified CtBP/BARS (5), $10 \mu\text{M}$ pCoA (6), CtBP/BARS and pCoA (7), $10 \mu\text{g ml}^{-1}$ ADP-ribosylated CtBP/BARS (8), ADP-ribosylated CtBP/BARS and pCoA (9). The fission index of untreated membranes (~ 5 ; stripes) represents background fission. Bars are mean $\pm$ s.d. of three independent experiments. **d**, Fission/fragmentation activity of partially purified and recombinant CtBP/BARS in the presence of pCoA. Golgi membranes were incubated for 15 min in $10 \mu\text{M}$ CoA and increasing concentrations of either partially purified or recombinant CtBP/BARS. Values are mean $\pm$ s.d. of three independent experiments.

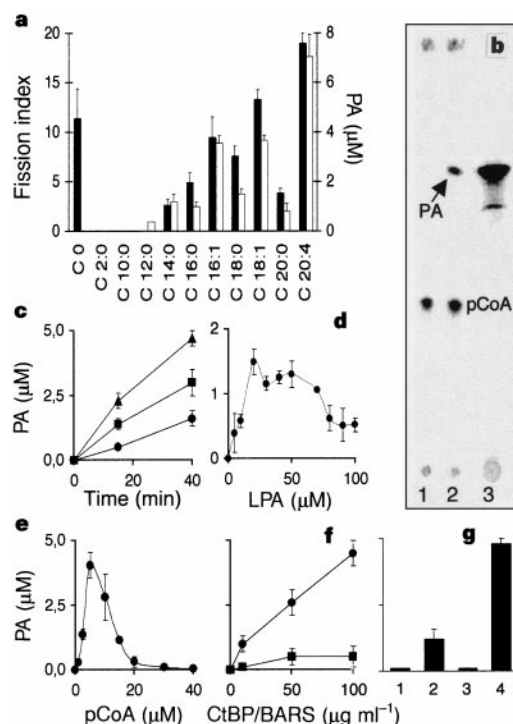


Figure 4 Acyl-CoAs support CtBP/BARS-dependent LPA acyltransferase and fission of Golgi membranes. **a**, Golgi membranes were incubated with different acyl-CoAs ($10 \mu\text{M}$) for 5 min in $10 \mu\text{g ml}^{-1}$ recombinant CtBP/BARS, and membrane fission (full bars) was quantified. LPA acyltransferase activity (open bars) was assessed by incubating recombinant CtBP/BARS for 40 min using [³H]C18:1 LPA ($30 \mu\text{M}$) and different unlabelled acyl-CoAs ($10 \mu\text{M}$) as substrates. Tritium on TLC plates was quantified using a linear analyser INB 384 (Inotech, Switzerland). Values are mean $\pm$ s.d. of three independent experiments. **b**, Synthesis of PA. [¹⁴C]pCoA ($10 \mu\text{M}$) was incubated with $30 \mu\text{M}$ LPA (C18:1-LPA) in the absence (lane 1) or presence (lane 2) of $50 \mu\text{g ml}^{-1}$ recombinant CtBP/BARS for 40 min. The [¹⁴C]PA standard is in lane 3. Lipids were analysed by TLC. Other lysolipids, headgroups or lipids were used at various concentrations (0.1 – $100 \mu\text{M}$; see Methods), as potential substrates for the transferase activity of recombinant CtBP/BARS. They were all inactive at all concentrations tested. **c–g**, Characterization of CtBP/BARS-dependent LPA acyltransferase activity. **c**, Relationship between CtBP/BARS concentration and yield of PA. $25 \mu\text{g ml}^{-1}$ (circles), $50 \mu\text{g ml}^{-1}$ (squares) or $100 \mu\text{g ml}^{-1}$ (triangles) of recombinant CtBP/BARS were incubated for 15 or 40 min with $5 \mu\text{M}$ pCoA and $30 \mu\text{M}$ LPA. **d, e**, Relationship between substrate concentrations and yield of PA. Recombinant CtBP/BARS ($50 \mu\text{g ml}^{-1}$) was incubated for 40 min (see Methods) with different concentrations of LPA (**d**) or pCoA (**e**). Values are mean $\pm$ s.d. of three independent experiments. **f**, Effect of ADP-ribosylation on CtBP/BARS-dependent LPA-acyltransferase activity. Increasing concentrations of recombinant CtBP/BARS either in the native (circles) or ADP-ribosylated (squares) forms were incubated for 40 min with $5 \mu\text{M}$ pCoA and $30 \mu\text{M}$ LPA. Values are mean $\pm$ s.d. of three independent experiments. **g**, Effect of incorporation of LPA into liposomes on CtBP/BARS acyltransferase activity. $50 \mu\text{g ml}^{-1}$ of either recombinant CtBP/BARS (2, 4) or GST (1, 3) were incubated with $5 \mu\text{M}$ pCoA for 15 min with $30 \mu\text{M}$ LPA resuspended in 25 mM HEPES pH 7.4 (1, 2), or included in Golgi-derived liposomes (3, 4). Values are mean $\pm$ s.d. of three independent experiments.

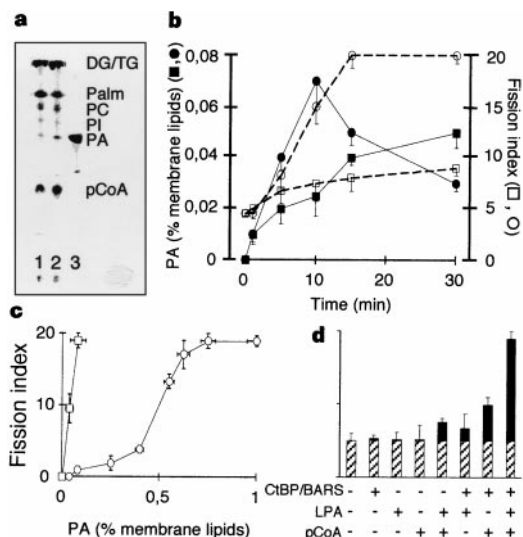


Figure 5 PA synthesis in Golgi membranes is associated with membrane fission/fragmentation. **a**, Golgi membranes were incubated with 10 μM [^{14}C]pCoA in the absence (lane 1) or presence (lane 2) of 20 $\mu\text{g ml}^{-1}$ recombinant CtBP/BARS for 10 min. The label was incorporated into the lipids indicated in the figure (identified by co-migration with unlabelled standards). PA was increased in the presence of CtBP/BARS. **b**, Time-course of PA formation and tubule fission/fragmentation in Golgi membranes. Membranes were incubated as described in **a** in the absence (squares) or presence (circles) of recombinant CtBP/BARS and processed for lipid analysis or negative-stain EM. PA levels, calculated as percentage of total membrane lipids, are shown on the left axis (filled symbols, continuous lines); the index of fission/fragmentation is shown on the right axis (open symbols, dashed lines). Values are mean $\pm$ s.d. of three independent experiments. **c**, Golgi membranes were incubated for 15 min in 10 $\mu\text{g ml}^{-1}$ recombinant CtBP/BARS and 5 or 10 μM [^{14}C]pCoA (squares), or in increasing concentrations of pCoA (1–10 μM) plus [^3H]C18:1 LPA (1–10 μM ; circles), and the newly formed PA was measured. Parallel unlabelled samples were obtained and processed for EM. The fission index is reported as a function of the PA levels in Golgi membranes. Values are means $\pm$ s.d. of three independent experiments; **d**, Golgi membranes were incubated for 15 min with 2 μM pCoA, 2 μM C18:1 LPA and 10 $\mu\text{g ml}^{-1}$ recombinant CtBP/BARS, either alone or in combination, as indicated in the graph. The fission index (see Methods) of untreated membranes (~ 5 ; stripes) represents background fragmentation. Values are mean $\pm$ s.d. of three independent experiments.

shown). Finally, the preference of CtBP/BARS acyltransferase for acyl-CoAs with different acyl chains was tested in incubations with [^3H]LPA (10–30 μM) to measure formation of PA. Short-chain ($C < 12$) acyl-CoAs were inactive; above C12, the reaction rate increased with chain length, and was maximum with unsaturated long-chain compounds such as oleoyl- (C18:1) and arachidonoyl- (C20:4) CoAs (Fig. 4a, white bars). The latter gave a reaction rate of 10–15 cycles min^{-1} per molecule of CtBP/BARS, using LPA embedded in liposomes and empirically optimized conditions (data not shown). This relatively low value might depend on the lack of activators, which instead might be present in Golgi membranes. The rank order of potency of the acyl-CoAs was the same as in the Golgi tubule fission assays (Fig. 4a), in agreement with the enzymatic and fission activities of CtBP/BARS being causally linked.

The transferase activity of CtBP/BARS, if it is functionally relevant, should cause changes in PA levels in Golgi membranes under the conditions used in the fission assay. We first characterized the enzymatic activity of the membranes themselves. Golgi membranes hydrolysed [^{14}C]pCoA to [^{14}C]palmitic acid and CoA, and incorporated the fatty acid into several major lipids. [^{14}C]PA was also formed, albeit in smaller amounts (Fig. 5a). Thus, Golgi membranes have significant acyltransferase activity. When CtBP/BARS was added to this background, the formation of PA increased threefold; this increase was transient, with a peak at 10–15 min. The

labelling of other lipids was unchanged (Fig. 5a, b), indicating that CtBP/BARS does not activate other PA-forming enzymes, such as phospholipase D (ref. 8), in Golgi membranes. We then compared the CtBP/BARS-induced increase in PA levels and the fission response in parallel experiments: the two processes occurred within the same time interval (Fig. 5b), indicating that they are associated. Finally, we examined whether the LPA acyltransferase reaction is sufficient and necessary for membrane fission. We exploited the endogenous LPA acyl-transferase activity of Golgi membranes. Membranes were incubated with LPA and pCoA (see legend to Fig. 5c), and PA synthesis and membrane fission were monitored. PA synthesis was proportional to the concentrations of LPA and pCoA used. Fission began to occur at levels of PA above 0.2% of the membrane lipids, and reached a maximum at 0.6% (Fig. 5c). The resulting fragments were similar to those induced by CtBP/BARS. Constricted fission intermediates were also observed. However, not only tubules but also Golgi cisternae underwent fragmentation, possibly because PA is synthesized in all of the Golgi components by the endogenous transferases. LPA without pCoA had no fissionogenic activity. Other lysolipids incubated with pCoA showed no sign of fissionogenic activity (data not shown). Thus, the acylation of LPA to PA is sufficient to cause membrane fission, even in the absence of CtBP/BARS. However, less PA (6–10-fold) was needed for CtBP/BARS-induced fission (Fig. 5c), possibly because CtBP/BARS acts only on tubules, whereas endogenous transferases act throughout the Golgi membranes.

We also wanted to establish whether, under suitable conditions, there was a requirement for LPA for fission. Golgi membranes were incubated with levels of pCoA and CtBP/BARS just above the threshold for inducing fission (Fig. 5d). Under these conditions, the endogenous LPA in Golgi membranes should be barely sufficient to drive the transferase activity of CtBP/BARS to the levels required for fission; therefore, the addition of LPA should enhance the reaction and trigger fission. Indeed, 2 μM LPA induced massive fission/fragmentation of the Golgi tubular networks. In the absence of CtBP/BARS, these levels of LPA or pCoA alone were inactive, and, combined, only slightly active (Fig. 5d). Other lysolipids had no effect. Finally, we tested the effects of an anti-BARS antibody (PEP9)². This treatment inhibited both CtBP/BARS-induced PA formation and membrane fission by 70–90%, similar to the effects of CtBP/BARS ADP-ribosylation (see Figs 3c and 4f). These results further confirm the causal link between the LPA acyltransferase and the fissionogenic activity of CtBP/BARS. In addition, they rule out the possibility that the transferase activity of recombinant CtBP/BARS is due to bacterial contaminants. Consistent with this conclusion, CtBP/BARS-expressing bacteria have significantly higher LPA acyltransferase activity than control cells (data not shown). At this point, it is also interesting to note that CtBP/BARS has a high structural similarity to the CtBP family of transcription regulators². Thus, CtBP/BARS might be a multifunctional protein¹⁰ and have an additional role in transcription. The possible dual role of CtBP/BARS has been discussed previously².

Altogether, our data show that the LPA acyltransferase activity of CtBP/BARS is an essential component of the mechanism by which this protein promotes fission in Golgi tubular networks. What is unique about LPA/PA? And what are the mechanisms by which the LPA acyltransferase reaction triggers fission? PA does have unique physical properties. It has a small headgroup with a high charge density, capable of acting both as a hydrogen donor and acceptor. Thus, a reduction in ionization of PA, for example upon binding divalent cations, results in the formation of PA-rich microdomains through intermolecular hydrogen bonding; the same probably holds for LPA¹¹. Moreover, the shape of PA can vary depending on the presence of calcium: its normal cylindrical shape becomes conical in millimolar calcium^{7,9}. In contrast, LPA has an inverted-cone shape⁷. Moreover, PA can be rapidly dephosphorylated to diacylglycerol, which is in itself strongly conical⁷. Diacylglycerol can

spontaneously flip-flop across the bilayer, and hence affect the composition of the luminal leaflet even when it is formed at the cytosolic surface⁷. Thus, a lipid microdomain containing LPA, PA and diacylglycerol rapidly interconverting has the potential to act as the lipid machinery driving, or facilitating, both the formation of highly curved fission intermediates and the overall process of fission, through coordinated changes in local membrane curvature. We feel that a more detailed model of lipid-driven fission is premature at this time. The lipid-fission machinery might be more complex and involve phospholipases, PA phosphatases, or flippases, as well as lipid-binding proteins, such as the phosphatidylinositol carrier protein implicated in fission in the Golgi¹².

The best characterized fission protein so far, dynamin, acts in concert with other proteins at the neck of endocytic vesicles^{13,14}. One dynamin isoform has been suggested to localize in the Golgi complex^{15,16}. Thus, an obvious issue is whether dynamin contributes to CtBP/BARS-dependent fission. Dynamin is very scarce in our Golgi preparations, and the structural effects of the two proteins on *in vitro* membranes are quite different^{13,14}, suggesting that the two machineries are unrelated. However, an element of commonality is emerging between the CtBP/BARS- and the dynamin-driven fission machineries: a dynamin-binding protein, SH3p4 (endophilin¹⁸), has been independently proposed to be required for the formation of endocytic vesicles¹⁷, and to have, like CtBP/BARS, LPA acyltransferase enzymatic activity. This suggests that dynamin, in addition to its proposed mechanochemical activity¹³, requires synergy with a lipid metabolic pathway, the same pathway that is activated by CtBP/BARS. Another potentially analogous situation is the fission of COPI-coated vesicles, for which a requirement for acyl-CoA was shown several years ago¹. These converging observations suggest that CtBP/BARS is a prototype of a new class of fission-inducing proteins active in lipid metabolism, and that acylation of LPA is a common ingredient in different membrane-fission events. □

Methods

Golgi membranes, CtBP/BARS, antibodies and cytosol

Golgi membranes¹⁹, partially purified rat brain CtBP/S¹, ADP-ribosylated CtBP/BARS¹, CtBP/BARS-depleted cytosol², anti-BARS antibodies² and rat brain cytosol²⁰ were prepared as described previously. Rat brain CtBP/BARS was enriched 900-fold². Further purification resulted in erratic activity. To obtain recombinant CtBP/BARS, the 1290-bp coding sequence of CtBP/BARS cDNA was amplified by PCR using specific primers, and cloned in ECORI- and NOTI-cleaved pGEX4T1 vector (Pharmacia Biotech). Glutathione S-transferase (GST) and GST-CtBP/BARS were expressed and purified according to GST Gene Fusion System (Pharmacia Biotech). GST-CtBP/BARS was completely pure, as judged by SDS-PAGE and silver staining. However, it was less active (mole by mole) than brain CtBP/BARS, consistent with the fact that only 10–30% of it could be ADP-ribosylated; our experience indicates that the ability to ADP-ribosylate CtBP/BARS tightly correlates with activity. When GST-CtBP/BARS was used, controls included GST. Golgi membranes (0.1 mg ml⁻¹), 70% pure by morphological and compositional (using markers of intracellular organelles) criteria, were incubated at 37 °C in 25 mM HEPES pH 7.4, 50 mM KCl, 200 mM sucrose, 1 mM ATP, 1 mM GTP, 5 mM DTT, 2.5 mM MgCl₂, 10 U ml⁻¹ creatine phosphokinase, 10 mM creatine phosphate (incubation buffer). Reactions were stopped with 0.5% or 2% glutaraldehyde, for assays using negative-stained and resin-embedded membranes, respectively.

Electron microscopy

For negative staining of whole-mount preparations, fixed samples were placed on a 1% formvar-coated grid for 5 min, and washed with 20 mM HEPES pH 6.8, covered with 10 mg ml⁻¹ BSA for 1 min, washed thrice with 20 mM HEPES pH 6.8. The contrast (NanoW 2018, Nanoprobes) was applied for 2 × 20 s. For resin-embedded membranes, fixed membranes were mixed with an equal volume of 2% BSA for 30 min, then centrifuged (10,000g for 30 min, to collect stacks) or ultracentrifuged (150,000g for 1 h, to collect also light membranes released from stacks), stained with OsO₄ and embedded as described²¹.

In negative-stain preparations, the fission/fragmenting effect of CtBP/BARS was quantified in a single-blind fashion. 30–100 Golgi stacks per grid were selected by systematic random sampling²². A 'fission index' was evaluated for each stack using four criteria: (1) disappearance of tubular-reticular structures and fenestrations; (2) number of vesicular structures (20 vesicles per stack was the control value in untreated membranes); (3) number of fission intermediates; (4) fraction of naked over tubule-decorated cisternae. Parameter 4 was replaced by 'disappearance of recognizable cisternae' in the case of experiments in which fission was induced in the absence of CtBP/BARS (Fig. 5c) because under those conditions not only tubules, but also cisternae, became fragmented.

These parameters varied with identical time courses under all conditions, except that the increase in number of constriction sites usually slightly preceded the other three events scored. Each criterion was given a score from 0 to 5 (for a total of 20). Four grids were evaluated for each experimental condition. For example, control Golgi membranes received a score of ~5 (low degree of fragmentation), membranes treated with CtBP/BARS-enriched cytosol for 30 min, a score of 18.

Assay of acyltransferase activity

GST-CtBP/BARS was incubated with [³H]LPA and unlabelled acyl-CoAs or, alternatively, with [¹⁴C]pCoA and LPA (or C14:0, C16:0 and C18:0 lysophosphatidylcholine, C16:0 and C18:0 lysophosphatidylethanolamine, lysophosphatidylinositol, lysophosphatidylserine, sphingosine 1-phosphate, glycerol, 3-glycerophosphate, glycerophosphocholine, glycerophosphoinositol, glycerophosphoethanolamine, dihydroxyacetone, phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, PA). In some experiments, LPA was first incorporated into Golgi-derived liposomes, prepared as described²³. The reaction was stopped by adding 1 volume of cold methanol/8 M HCl (1500:1). Lipids were extracted in 2 volumes of chloroform/methanol (2:1), loaded on an oxalate-pretreated TLC plate (Kieselgel 60, Merck) and eluted with chloroform/methanol/20% NH₄OH/H₂O (54:42:2.9:1). Radiolabelled spots were quantified by autoradiography (Instant Imager, Canberra Packard).

Received 16 June; accepted 12 October 1999.

- Ostermann, J. *et al.* Stepwise assembly of functionally active transport vesicles. *Cell* **75**, 1015–1025 (1993).
- Spanò, S. *et al.* Molecular cloning and functional characterization of brefeldin A-ADP-ribosylated substrate. A novel protein involved in the maintenance of the Golgi structure. *J. Biol. Chem.* **274**, 17705–17710 (1999).
- Weidman, P., Roth, R. & Heuser, J. Golgi membrane dynamics imaged by freeze-etch electron microscopy: views of different membrane coatings involved in tubulation versus vesiculation. *Cell* **75**, 123–133 (1993).
- Rambourg, A. & Clermont, Y. in *The Golgi Apparatus* (eds Berger, E. G. & Roth, J.) 37–60 (Birkhäuser, Basel, 1997).
- Nakamura, N., Lowe, M., Levine, T. P., Rabouille, C. & Warren, G. The vesicle docking protein p115 binds GM130, a *cis*-Golgi matrix protein, in a mitotically regulated manner. *Cell* **89**, 45–55 (1997).
- Jones, S. M. & Howell, K. E. Phosphatidylinositol 3-kinase is required for the formation of constitutive transport vesicles from the TGN. *J. Cell Biol.* **139**, 339–349 (1997).
- Cullis, P. P., Fenske, D. B. & Hope, M. in *Biochemistry of Lipids, Lipoproteins and Membranes* (eds Vance, D. E. & Vance, J. E.) 1–32 (Elsevier Science, Amsterdam, 1996).
- Kent, C. Eukaryotic phospholipid biosynthesis. *Annu. Rev. Biochem.* **64**, 315–343 (1995).
- Bar-Tana, J. & Shapiro, B. Long-chain fatty acyl-CoA synthetase from rat liver microsomes. EC 6.2.1.3 fatty acyl-CoA ligase (ATP). *Methods Enzymol.* **35**, 117–122 (1975).
- Jeffery, C. J. Moonlighting proteins. *Trends Biochem. Sci.* **24**, 8–11 (1999).
- Boggs, J.-M. Lipid intermolecular hydrogen bonding: influence on structural organization and membrane function. *Biochim. Biophys. Acta* **906**, 353–404 (1987).
- Simon, J.-P. *et al.* An essential role for the phosphatidylinositol transfer protein in the scission of coatomer-coated vesicles from the trans-Golgi network. *Proc. Natl Acad. Sci. USA* **95**, 11181–11186 (1998).
- Kelly, R. B. New twists for dynamin. *Nature Cell Biol.* **1**, E8–E9 (1999).
- Cremona, O. & De Camilli, P. Synaptic vesicle endocytosis. *Curr. Opin. Neurobiol.* **3**, 323–330 (1997).
- Cao, H., Garcia, F. & McNiven, M. A. Different distribution of dynamin isoforms in mammalian cells. *Mol. Biol. Cell* **9**, 2595–2609 (1998).
- Altschuler, Y. *et al.* Redundant and distinct functions for dynamin-1 and dynamin-2 isoforms. *J. Cell Biol.* **143**, 1871–1881 (1998).
- Schmidt, A. *et al.* Endophilin I mediates synaptic vesicle formation by transfer of arachidonate to lysophosphatidic acid. *Nature* **401**, 133–141 (1999).
- Ringstad, N., Nemoto, Y. & De Camilli, P. The SH3p4/SH3p8/SH3p13 protein family: binding partners for synaptotagmin and dynamin via a Grb2-like Src homology 3 domain. *Proc. Natl Acad. Sci. USA* **94**, 8569–8574 (1997).
- Cluett, E. B. & Brown, W. J. Adhesion of Golgi cisternae by proteinaceous interactions: intercisternal bridges as putative adhesive structures. *J. Cell Sci.* **103**, 773–784 (1992).
- Malhotra, V., Serafini, T., Orci, L., Sheperd, J. C. & Rothman, J. E. Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell* **58**, 329–336 (1989).
- Buccione, R. *et al.* Regulation of constitutive exocytic transport by membrane receptors. A biochemical and morphometric study. *J. Biol. Chem.* **271**, 3523–3533 (1996).
- Lucocq, J. Unbiased 3-D quantitation of ultrastructure in cell biology. *Trends Cell Biol.* **3**, 354–358 (1993).
- Takei, K. *et al.* Generation of coated intermediates of clathrin-mediated endocytosis on protein-free liposomes. *Cell* **94**, 131–141 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank A. Mironov Jr, G. Beznoussenko and R. Polishchuk for help with the electron microscopy; V. Malhotra for helpful discussions; C. P. Berrie for critical reading of the text; C. Iurisci for technical assistance; J. Füllekrug, K. Moremen, M. Zerial, G. Banting, A. Hubbard, E. Sztul, H. P. Hauri and V. Malhotra for their generous gifts of antibodies. This work was supported in part by grants from Telethon-Italy, AIRC, and CNR-Progetti Singoli e Finalizzato "Biotecnologie". S.S. and R.M. are recipients of FIRC fellowships.

Correspondence and requests for materials should be addressed to D.C. (e-mail: corda@cmns.mnegr.it) or A.L. (e-mail: luini@cmns.mnegr.it).

Crystal structures of the membrane-binding C2 domain of human coagulation factor V

Sandra Macedo-Ribeiro*, Wolfram Bode*, Robert Huber*, Mary Ann Quinn-Allen†, Suhng Wook Kim†, Thomas L. Ortel†, Gleb P. Bourenkov‡, Hans D. Bartunik‡, Milton T. Stubbs§, William H. Kane† & Pablo Fuentes-Prior*

* Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, 82152 Martinsried, Germany

† Division of Hematology, Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710, USA

‡ Max-Planck Research Unit for Structural Molecular Biology, Protein Dynamics Group, MPG-ASMB c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

§ Institut für Pharmazeutische Chemie, Philipps Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

Rapid and controlled clot formation is achieved through sequential activation of circulating serine proteinase precursors on phosphatidylserine-rich procoagulant membranes of activated platelets and endothelial cells¹. The homologous complexes Xase and prothrombinase, each consisting of an active proteinase and a non-enzymatic cofactor, perform critical steps within this coagulation cascade. The activated cofactors VIIIa and Va, highly specific for their cognate proteinases, are each derived from precursors with the same A1-A2-B-A3-C1-C2 architecture². Membrane binding is mediated by the C2 domains of both cofactors. Here we report two crystal structures of the C2 domain of human factor Va. The conserved β -barrel framework provides a scaffold for three protruding loops, one of which adopts markedly different conformations in the two crystal forms. We propose a mechanism of calcium-independent, stereospecific binding of factors Va and VIIIa to phospholipid membranes^{3,4}, on the basis of (1) immersion of hydrophobic residues at the apices of these loops in the apolar membrane core; (2) specific interactions with phosphatidylserine head groups in the groove enclosed by these

loops; and (3) favourable electrostatic contacts of basic side chains with negatively charged membrane phosphate groups.

Although the calcium-mediated membrane binding of vitamin-K-dependent coagulation factors has been studied intensively^{5,6}, the calcium-independent phosphatidyl-L-serine (P₁S)-specific interaction of cofactors Va⁴ and VIIIa⁷ is poorly understood. Upon membrane-mediated complex formation, activated factor V (FVa) composed of the heavy A1-A2 chain and the light A3-C1-C2 chain² enhances the ability of factor Xa to generate α -thrombin from prothrombin by five orders of magnitude⁸. The C2 domain of FVa (FVa-C2) is essential for binding to P₁S-containing membranes and thus for procoagulant activity³. FVa-C2 is homologous to the discoidin family of adhesion proteins, but is not related to the synaptotagmin-like C2 domains.

We have solved the structure of FVa-C2 in a monomeric and a dimeric crystal form. FVa-C2 exhibits a distorted jelly-roll β -barrel motif, consisting of eight major antiparallel strands arranged in two β -sheets of five and three strands packed against one another (Fig. 1). This barrel is closed at its top and bottom by three and two straight segments, respectively, giving it a spherical shape with a flattened upper surface. Whereas the upper part of the barrel surface contains several notable salt bridges (for example, Asp 61-Arg 134) (C2 residue numbering; Fig. 2), its lower part is characterized by a preponderance of basic residues (Fig. 1a). Both amino- and carboxy-terminal segments are clamped together by the single disulphide bridge Cys 1-Cys 156 located at the edge of the upper flat surface.

Three adjacent loops (Ser 21-Trp 31, Asn 39-Asn 45 and Gly 75-Tyr 84) protrude like spikes from the bottom of the β -barrel in monomeric FVa-C2 (Fig. 1). Spikes 1 and 3 consist of separate β -hairpin structures, whereas spike 2 is a wider irregular loop. The three spikes are linked to one another and to three shorter loops by an intricate hydrogen-bonding network which extends to residues at the bottom of the β -barrel. At the apex of spike 1, the indole moieties of two consecutive tryptophan residues (Trp 26 and Trp 27) are fully exposed to solvent. Spike 3 also harbours a hydrophobic residue at its apex (Leu 79), whereas spike 2 is capped by the Arg 43 side chain (Fig. 1a). The three circularly arranged spikes create a deep groove lined by both hydrophobic (Trp 31, Met 83) and polar (Gln 48, Ser 78) side chains (Fig. 1a). We refer to this structure as the 'open form' of FVa-C2.

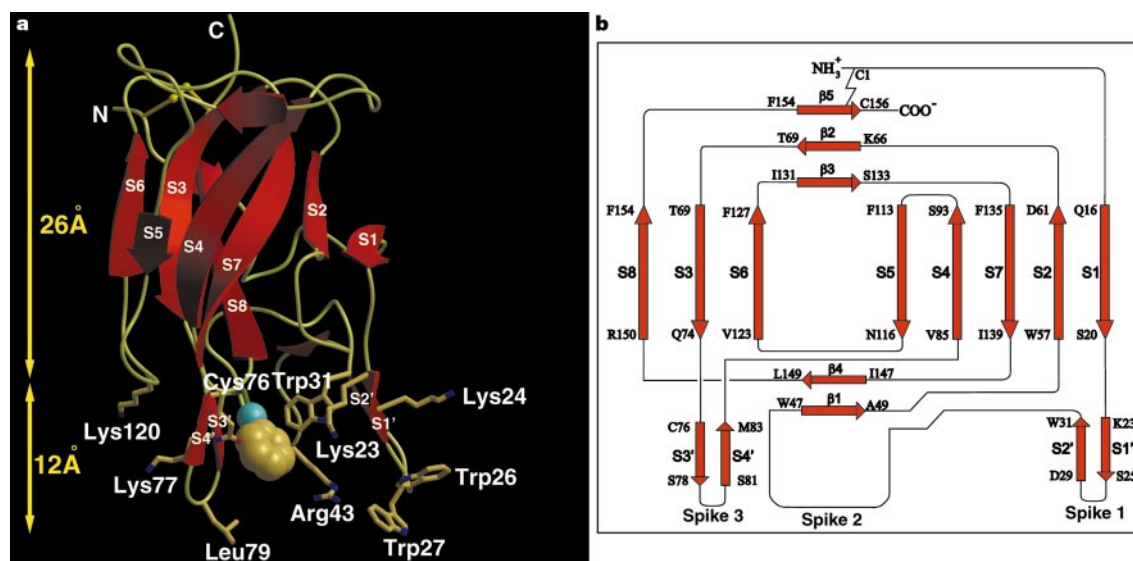


Figure 1 Structure of FVa-C2. **a**, Ribbon plot of the 'open' crystal form in the reference orientation, highlighting the major secondary structure elements. Several noticeable side chains and the Cys 1-Cys 156 disulphide bridge are displayed. The phenyl mercury

molecule (space-filling model) used for structure determination is shown covalently attached to Cys 76. **b**, Topology diagram showing the arrangement of secondary-structure elements in FVa-C2.

In the second, dimeric crystal form of FVa-C2, both molecules are quasisymmetrically packed through the free edges of their S6 strands. The only significant difference between the two crystallographically independent molecules is in the Leu 104–Val 109 loop protruding from the upper flat surface. In the spike region, however, both molecules differ significantly from the open form of FVa-C2 (Fig. 3a). The first and third spikes are tilted towards the interior of the groove, with spike 1 additionally deformed because of a strong twist around Gly 28. This concerted tilting–twisting movement not only results in main-chain atom shifts of up to 7 Å, but also leads to a 12-Å displacement of the Trp 27 indole moiety, locating it close to the apices of the two other spikes (Fig. 3a,b). The side chain of Trp 26 moves in concert with Trp 27 and packs against it, leaving a small entrance to the shallow interspike groove, in particular to the critical Gln 48 carboxamide (see below). The side chains of Trp 26, Trp 27 and Leu 79 form a hydrophobic ridge (Fig. 3b), which covers the groove observed in the open form (Fig. 3a, c). Therefore, we call this structure the ‘closed form’ of FVa-C2.

The region delimited by the three spikes can be assigned to the primary membrane-binding site of the C2 domain (and therefore of factors Va and VIIIa)^{9,10}. The monoclonal antibody HV-1, which maps to spike 1 (ref. 11), interferes with P₁S-specific binding and abolishes procoagulant activity^{12,13}. Glycosylation of Asn 144 in the neighbouring S7–S8 loop (Fig. 1a, 3a) interferes with membrane binding of FVa^{14,15}. Substitution of alanine residues for Trp 26 and Trp 27 blocks binding of FVa to immobilized P₁S and prevents expression of procoagulant activity in the presence of limiting concentrations of membranes¹¹. The N-terminal domain of a fungal galactose oxidase¹⁶ (N-GOase) exhibits structural similarity to FVa-C2 (refs 9, 10, 17) (Fig. 4), both modules representing the first members of their respective subfamilies¹⁷ for which three-dimensional structures have been reported. N-GOase contains a secondary galactose-binding site in a region equivalent to the interspike groove of FVa-C2 (Fig. 4a), which has been predicted to anchor the enzyme to plant cell walls. This may indicate a common binding role for the spike region (Fig. 4b).

We propose that the open conformation of FVa-C2 corresponds to the membrane-bound molecular state, whereas the closed form represents the major species circulating in plasma. The markedly smaller (370 Å compared with 520 Å²) hydrophobic surface exposed by residues Trp 26, Trp 27 and Leu 79 indicates that the closed form may be favoured in aqueous solution. As typical epitopes are flat,

convex, molecular surfaces, antibodies directed against the spike region (HV-1)^{11,13} probably recognize the closed form (see Fig. 3). Mercury compounds failed to bind to Cys 76 in the closed crystal form, and the two free thiol groups in the FVa light chain (Fig. 2) are oxidized only slowly under native conditions¹⁸, providing further support for this hypothesis.

In the open crystal form, the exposed indoles of Trp 26 and Trp 27 and the aliphatic side chain of Leu 79 constitute particularly good candidates for immersion in the apolar membrane core. This is reminiscent of the anticoagulant protein annexin V, in which the indole side chain of the unique tryptophan residue becomes exposed upon calcium binding to allow membrane insertion¹⁹. The interior of the groove is lined with polar side chains that are ideally arranged for interactions with head groups of the procoagulant lipid P₁S. Spikes 1 and 2 enclose the most obvious putative binding site (Fig. 5a). Within this ‘P₁S-specificity pocket’ the serine carboxylate of P₁S could accept hydrogen bonds from the carboxamide nitrogen of Gln 48 and the hydroxyl group of Ser 78, whereas its amino group is within hydrogen-bonding distance of the Gln 48 side-chain oxygen. The ϵ -amino group of Lys 23 would compensate the negative charge of the P₁S phosphate group. These residues are strictly conserved within the FVa-C2 domains (Fig. 2). Highly directional interactions inside this pocket would be strengthened by the hydrophobic environment made by the adjacent fatty-acid tails and conserved hydrophobic residues (Trp 27, Trp 31, Met 83, Leu 79). Another specific P₁S-binding site can be localized between the second and the third spike lined by the side chains of Gln 41, Asn 45 and Arg 150 and the carbonyls of Gly 42 and Lys 77. This second, strongly basic binding site (Fig. 3b, c) would be more accessible to acidic phospholipids.

The suitably placed polar head groups of FVa-C2 allow discrimination of P₁S over other acidic phospholipids²⁰, by stereospecific recognition of the P₁S head group⁴. Notably, the apparent dissociation constant (K_d) for membrane binding of FVa increases 25-fold upon replacing P₁S- by phosphatidyl-D-serine (P_DS)-containing membranes⁴. These stereospecific interactions explain why positively charged vesicles support normal prothrombinase activity *in vitro* if they contain P₁S, but not P_DS or phosphatidyl- β -lactate^{4,21}. Furthermore, they explain the importance of both the carboxyl²⁰ and the amino groups of P₁S²¹ for formation of a fully active prothrombinase complex.

We propose the following mechanism for the stereospecific

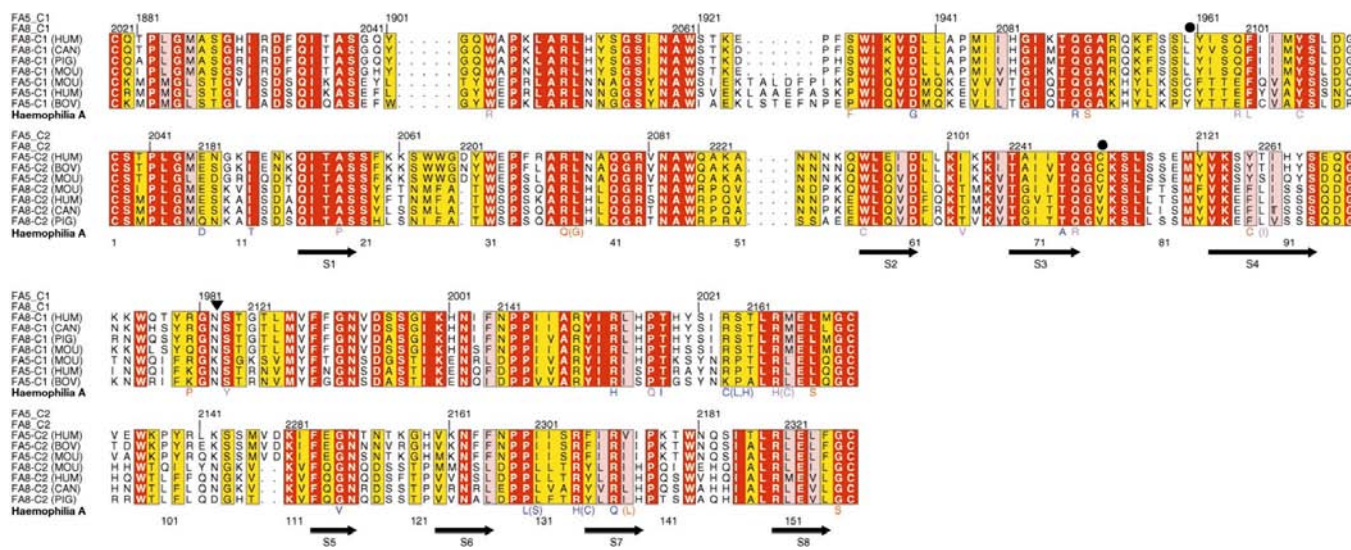


Figure 2 Sequence alignment of the C1 and C2 domains of human coagulation cofactors. Numbers for the full-length human cofactors are given above the alignment; the FVa-C2 numbering used in this paper is indicated below. Red, strictly conserved residues within C1 or C2 domains; pink, substitutions within the hydrophobic core; yellow, other

conservative substitutions. Glycosylation sites (closed triangles) and free cysteines (closed circles) are indicated. Missense mutations deposited in the haemophilia A database (URL: <http://europium.mrc.rpms.ac.uk>) are given in red (severe cases), pink (moderate) or blue (mild). Notice that most mutations affect conserved residues.

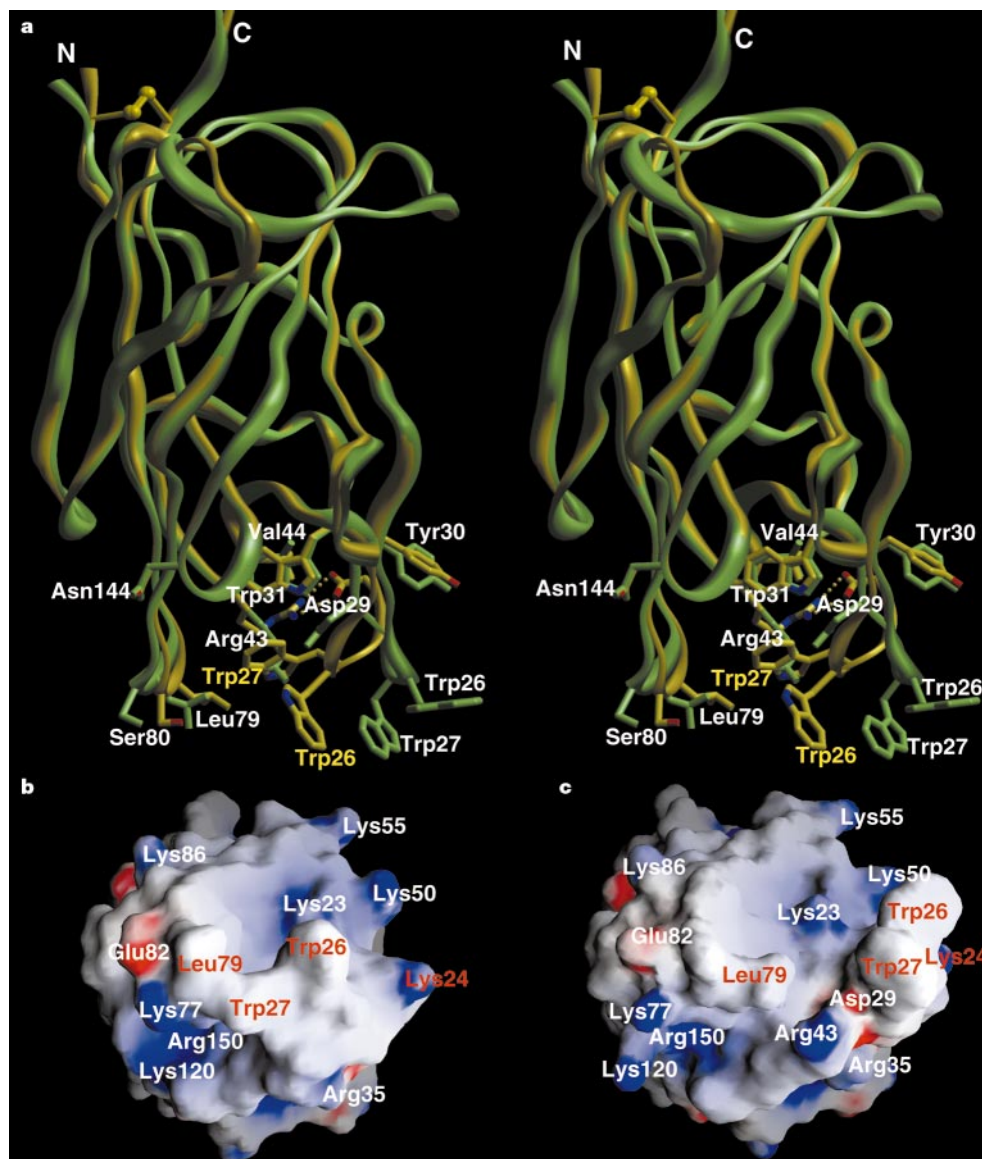


Figure 3 Conformational flexibility in the spike region. **a**, Stereo view of the FVa-C2 structures, superimposed in the same orientation as in Fig. 1. The open form is shown in green. The closed form is displayed in yellow, with several side chains colour-coded (carbon, yellow; oxygen, red; nitrogen, blue). **b, c**, GRASP³⁰ surface representations of the

electrostatic potential (red, negative; blue, positive) of FVa-C2 in the closed (**b**) and in the open (**c**) crystal form. The view is from the membrane side, that is, perpendicular to the reference orientation.

recognition and cooperative association with P₁S-rich membranes. (1) FVa, with the C2 domain in its closed form (Fig. 3b), approaches acidic membrane sites guided by favourable protein–membrane electrostatic interactions. (2) One or two P₁S molecules bind to anchoring points (such as Arg 150 and/or Gln 48) in the small groove of the closed form, triggering a widening of the groove and conversion to the open form (Fig. 3c). (3) A P₁S molecule occupies the opened specificity pocket (Fig. 5a). (4) The now unfolded hydrophobic spikes perforate the polar membrane surface, immersing the side chains of Trp 26, Trp 27 and Leu 79 in the apolar membrane core. (5) The bottom of the barrel contacts the negatively charged phosphate head groups of the membrane through favourable ionic interactions with basic residues located in the spikes and neighbouring loops (Figs 1a and 3), explaining the general preference of FVa for acidic phospholipid membranes^{6,20}. Because the spikes and neighbouring loops are highly interconnected, and several contiguous P₁S molecules would be necessary for membrane association, spike unfolding and membrane insertion would be

highly cooperative, assuring that FVa-C2 binds only to cell membranes with P₁S concentrations exceeding a critical threshold value. This feature would selectively target FVa to procoagulant, P₁S-rich surfaces.

Binding of FVa to a few acidic, lipid-specific sites results in substantial protein conformational changes²². Experimental evidence indicates that the A3 domain interacts with phosphatidylcholine-rich regions²³, and additional hydrophobic contributions of the heavy chain have been reported¹⁵. We are tempted to speculate that further conformational changes follow membrane association of FVa-C2. In particular, the C2 domain may rotate to bring not only spike 1, but also the flat and extended tryptophan-rich surface covering the front side of the β -barrel (Fig. 5b) in contact with the phospholipid membrane.

The proposed membrane immersion of the hydrophobic C2 spikes in the apolar membrane core resembles that postulated for γ -carboxyglutamic acid (Gla)-rich domains of vitamin-K-dependent coagulation factors⁵. However, whereas the hydrophobic

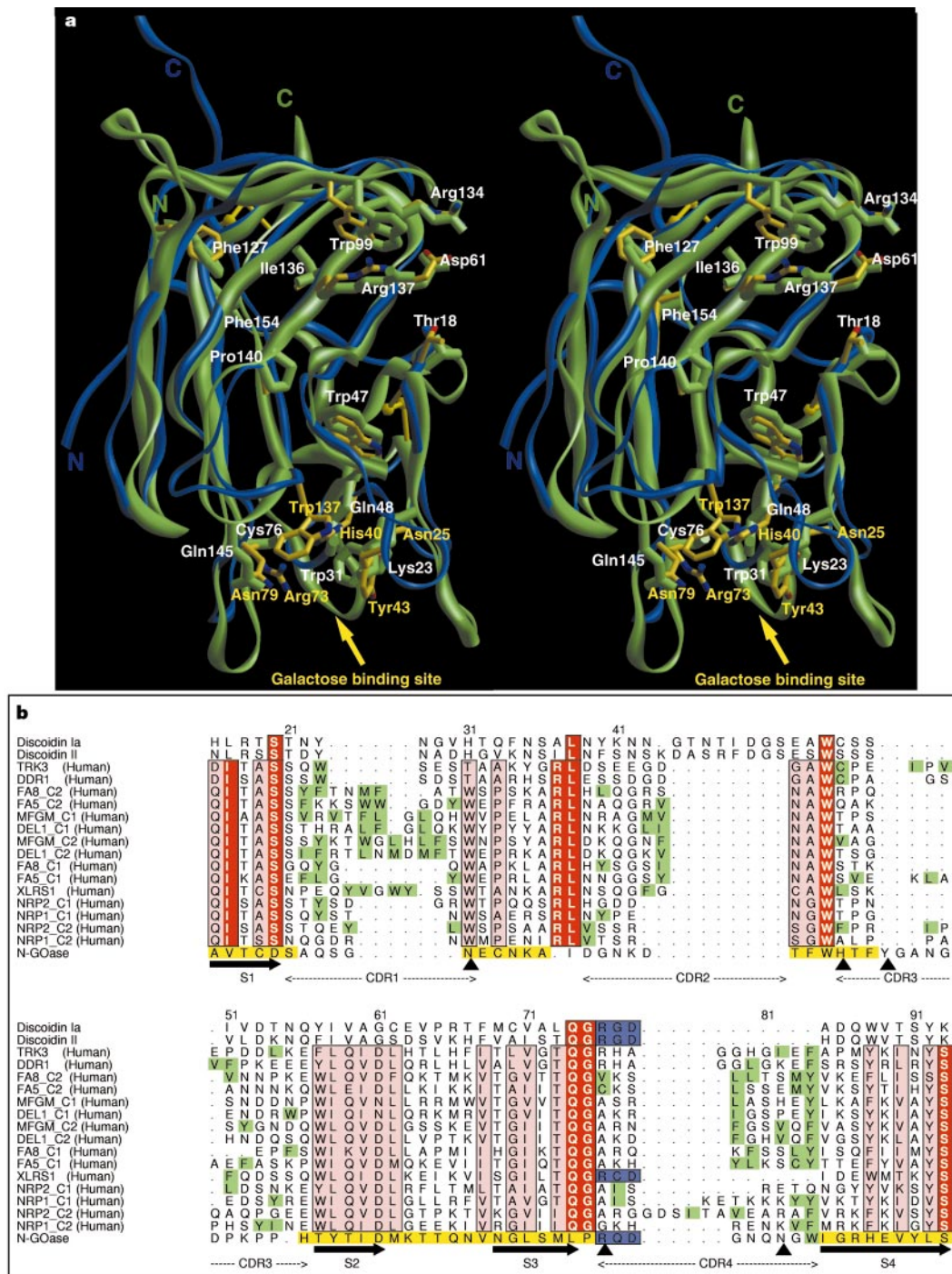


Figure 4 Variability within the discoidin family. **a**, Superposition of the crystal structures of FVa-C2 (green, open form) and N-GOase¹⁶ (blue). The side chains of several conserved residues in FVa-C2 (green) and in N-GOase (colour coded) are displayed. FVa-C2 side chains are labelled in white, N-GOase side chains involved in D-galactose binding in yellow. **b**, Sequence alignment of several representative discoidin domains. Numbers refer to the FVa-C2 sequence. Variable loops have been named CDR1 to CDR4, in analogy

to the immunoglobulin nomenclature. Topologically equivalent residues between FVa-C2 and N-GOase are shown in yellow; closed triangles indicate residues involved in galactose binding. The cell-adhesion motif RGD in discoidins Ia and II and related motifs are shown in blue. Hydrophobic/aromatic residues within the CDRs of several human proteins are shown in green; they are probably functionally relevant for other members of this family.

residues of structurally unrelated Gla domains are probably exposed at plasma calcium concentrations, a critical but low number of P₁S molecules are necessary to promote the subtle calcium-independent conformational changes¹⁸ necessary for FVa binding. The cooperative nature of this process presumably precludes binding of isolated P₁S head-group analogues to FVa-C2. Attempts to soak compounds such as O-phospho-L-serine and L-α-glycerophosphorylserine into

crystals of the open form showed stabilization only at the apex of spike 3. Notably, the phenyl moiety of PheH_g is poorly defined in XTAL1-PheH_g (Fig. 1a), despite covalent attachment to S_γ of Cys 76.

The structures of FVa-C2 constitute appropriate templates for modelling the C1 and C2 domains of FVa and FVIIIa, whose overall folds are certainly similar (Fig. 2). Several residues essential for

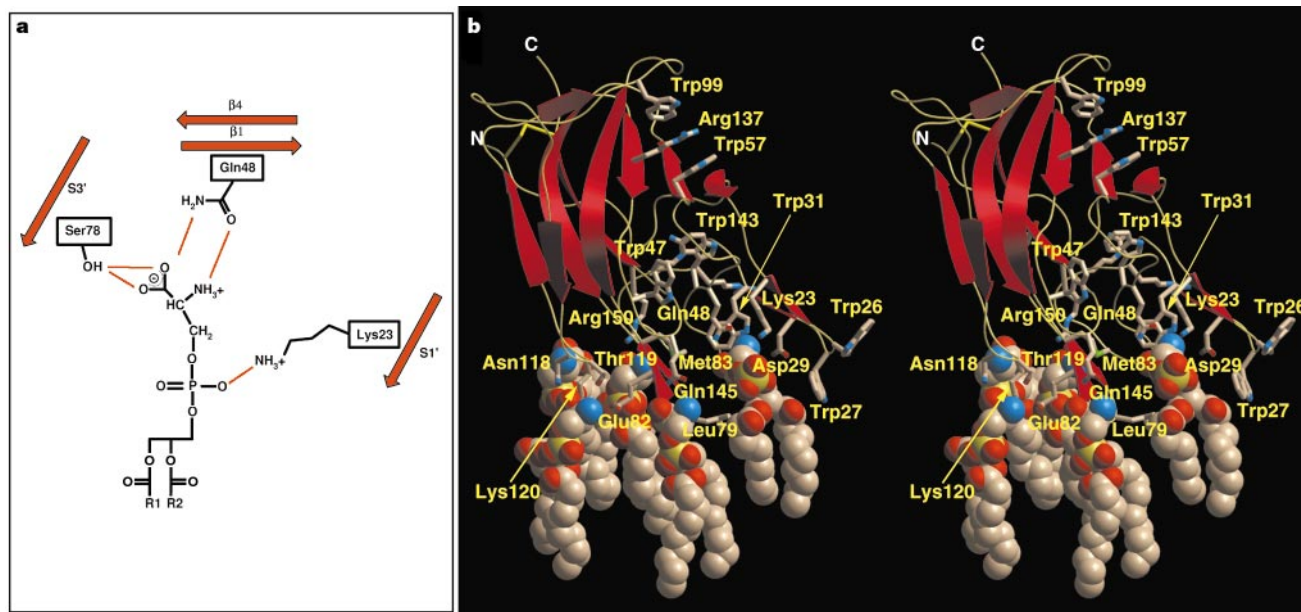


Figure 5 Structural basis of the stereospecific recognition of P_1S by FVa-C2. **a**, Predicted hydrogen-bond interactions (dashed red lines) inside the major P_1S -binding site. **b**, Stereo view of the 'open', membrane-competent form of FVa-C2 with five P_1S molecules bound

correct folding¹¹, such as Arg 37, Glu 152, the surface-located Asp 61–Arg 134 salt bridge or the characteristic Trp 57–Arg 137–Trp 99 triplet (Fig. 5b), are strictly conserved. Significant structural differences are expected for the spike regions and are probably functionally important. A five-residue deletion in both FVa-C1 and FVIIIa-C1 (Fig. 2) eliminates the critical first spike. Replacements such as Lys24Gly, Gln48Ser, Ser78His/Gln or Lys120Ser will either flatten the interspike groove or eliminate appropriate partners for interactions with acidic phospholipids³. However, the conservation of several aromatic/hydrophobic residues in the spike region of FVa-C1 (for example, Try 30, Tyr 44, Tyr 79, Tyr 144; see Fig. 2) may indicate its involvement in membrane binding, possibly through interactions with neutral membrane surfaces.

In the human FVIIIa-C2 model, the exposed Met 26 and Phe 27 that are equivalent to the two tryptophan residues of FVa-C2 (Figs 2, 4b) are likely candidates for membrane insertion. The substitutions Gln48Arg and Ser146Gln would reduce the size of the P_1S -specificity pocket to about one-quarter that of FVa-C2. Replacement of several lysine residues by uncharged/hydrophobic residues (Fig. 2) renders the spike region of FVIIIa-C2 more apolar and therefore less appropriate for interactions with negatively charged membranes. Accordingly, FVIIIa requires higher P_1S concentrations for membrane binding than FVa *in vitro*²⁴. The strict conservation of FVIIIa residues Gln 41, Asn 45 and Arg 150 (FVa-C2 residue numbering; Fig. 2) supports their predicted interaction with a P_1S head group, in line with experimental evidence⁷. Arg 48, which replaces the prominent FVa-C2 residue Gln 48, could specifically anchor a second P_1S molecule by forming a salt bridge with its free serine carboxylate.

These crystal structures provide insight into two aspects of the membrane association of factor Va, dynamics and complementarity. Instead of docking to procoagulant membranes through a rigid, pre-formed binding surface, the C2 domain of factor Va attains a conformation competent for favourable hydrophobic and electrostatic interactions with the phospholipid membrane only upon highly cooperative specific binding of P_1S . □

Methods

Purification and crystallization

A fragment of human factor V corresponding to residues Gly 2,037–Tyr 2,196 of the

to the proposed anchoring points in the spike region. Relevant side chains are labelled; bound phospholipid molecules are shown as space-filling models.

mature protein, preceded by an N-terminal sequence of Ala-Gln-Trp-Tyr-Gln, was produced using a baculovirus-driven insect cell system and purified as described¹³. After dialysis against 5 mM MOPS, pH 8.0, the sample was concentrated to ~6 mg ml⁻¹ and crystallized by vapour diffusion. Two different crystal forms were obtained. XTAL1 (space group $P2_12_12_1$, cell constants $a = 46.89$ Å, $b = 53.33$ Å, $c = 68.52$ Å) grows at 6 °C from 0.1 M Tris, 1.4 M ammonium sulphate, pH 10.2; XTAL2 (space group $P2_12_12_1$, $a = 86.52$ Å, $b = 70.54$ Å, $c = 60.58$ Å) at 20 °C from 0.1 M potassium phosphate, 0.7 M ammonium sulphate, pH 9.5–10.0. For collection of the multiple-wavelength anomalous-diffraction (MAD) data, XTAL1 crystals were soaked in mother liquor containing 5 mM phenylmercury (XTAL1-PheHg) and then transferred to a cryo-buffer (25–30% (v/v) glycerol, 1.4 M ammonium sulphate).

Data collection

MAD data from a single crystal (XTAL1-PheHg) were collected on the MPG beamline BW6 at DESY (Hamburg, Germany) using a MAR-CCD detector. Bijvoet pairs to 1.9 Å resolution were collected at the LIII absorption edge ($\lambda_1 = 1.0047$ Å, f''), at the inflection point ($\lambda_2 = 1.0100$ Å, maximal $\Delta f'$) and at a remote wavelength ($\lambda_0 = 1.05$ Å) in two continuous 90° sweeps. X-ray diffraction data for both XTAL1 (1.87 Å) and XTAL2 (2.40 Å) were collected in house using a Mar Research image plate system installed on a RIGAKU rotating anode generator. The diffraction images were evaluated with MOSFLM²⁵, and scaled and reduced with programs from the CCP4 program suite²⁶. XTAL1 diffraction data ($R_{\text{merge}} = 6.1\%$) are 94.0% complete (86.8% in the highest-resolution shell); the corresponding values for XTAL2 ($R_{\text{merge}} = 9.4\%$) are 96.8% and 90.7%, respectively.

Phasing and refinement

Data sets collected at wavelengths λ_1 and λ_2 were scaled to the λ_0 data using the program SCALEIT²⁶. The mercury site was determined with RSPS²⁷ from an anomalous difference Patterson map calculated using data with the highest anomalous differences. The maximum-likelihood program SHARP (URL: <http://lagrange.mrc-lmb.cam.ac.uk/sharp/SharpHome.phtml>) was used for heavy-atom refinement and phasing. Experimental phases were estimated at 1.9 Å resolution, giving a final figure of merit of 0.875 after iterative solvent flattening as implemented in SHARP. The straightforward interpretable density allowed building of an almost complete protein model with MAIN (URL: <http://www.bmb.jys.si/doc/>), with the exception of a few residues within some solvent-exposed loops. The structure was refined with X-PLOR²⁸ (1.0 σ cut-off) to an R -factor of 19.6% ($R_{\text{free}} = 25.8\%$) using 12,266 independent reflections of the λ_0 data set in the 8.0–1.9-Å resolution range. This model (XTAL-PheHg) was used for refinement (1.0 σ cut-off) against the XTAL1 data set, yielding a final model with an R -factor of 20.1% ($R_{\text{free}} = 24.2\%$) for 13,591 independent reflections in the 8.0–1.87-Å resolution range. The XTAL2 structure was solved by molecular replacement with AMoRe²⁹ using a XTAL1 truncated model, and crystallographically refined (1.0 σ cut-off) to an R -factor of 20.5% ($R_{\text{free}} = 25.7\%$) for 14,119 independent reflections and 8.0–2.40-Å data. In XTAL1, the polypeptide chain is defined except for residues Leu 79 and Ser 80, whereas the XTAL2 chain is fully defined. All non-glycine residues fall into allowed or additionally allowed regions.

Structural comparisons and modelling

Structural comparisons were made after least-squares superposition using the N-GOase coordinates as stored in the Protein Data Bank (PDB; accession code 1gog). Homology modelling was done using MAIN, with only a few manual interventions to avoid steric clashes. A P₁S model was constructed using the coordinates of L- α -glycerophosphorylserine in its annexin V complex¹⁹ (PDB code 1a8a), energy-minimized with X-PLOR²⁸ and manually fitted to the FVa-C2 structure.

Received 27 July; accepted 14 September 1999.

1. Davie, E. W., Fujikawa, K. & Kisiel, W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* **30**, 10363–10370 (1991).
2. Kane, W. H. & Davie, E. W. Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders. *Blood* **71**, 539–555 (1988).
3. Ortel, T. L., Devore-Carter, D., Quinn-Allen, M. A. & Kane, W. H. Deletion analysis of recombinant human factor V. Evidence for a phosphatidylserine binding site in the second C-type domain. *J. Biol. Chem.* **267**, 4189–4198 (1992).
4. Comfurius, P., Smeets, E. F., Willems, G. M., Bevers, E. M. & Zwaal, R. F. A. Assembly of the prothrombinase complex on lipid vesicles depends on the stereochemical configuration of the polar headgroup of phosphatidylserine. *Biochemistry* **33**, 10319–10324 (1994).
5. Stenflo, J. Contributions of Gla and EGF-like domains to the function of vitamin K-dependent coagulation factors. *Crit. Rev. Eukaryot. Gene Expression* **9**, 59–88 (1999).
6. Zwaal, R. F. A., Comfurius, P. & Bevers, E. M. Lipid–protein interactions in blood coagulation. *Biochim. Biophys. Acta* **1376**, 433–453 (1998).
7. Gilbert, G. E. & Drinkwater, D. Specific membrane binding of factor VIII is mediated by O-phospho-L-serine, a moiety of phosphatidylserine. *Biochemistry* **32**, 9577–9585 (1993).
8. Nesheim, M. E., Taswell, J. B. & Mann, K. G. The contribution of bovine factor V and factor Va to the activity of prothrombinase. *J. Biol. Chem.* **254**, 10952–10962 (1979).
9. Villoutreix, B. O., Bucher, P., Hofmann, K., Baumgartner, S. & Dahlbäck, B. Molecular models for the two discoidin domains of human blood coagulation factor V. *J. Mol. Modelling* **4**, 268–275 (1998).
10. Pellequer, J. L., Gale, A. J., Griffin, J. H. & Getzoff, E. D. Homology models of the C domains of blood coagulation factors V and VIII—a proposed membrane binding mode for FV and FVIII C2 domains. *Blood Cells, Molecules and Diseases* **24**, 448–461 (1998).
11. Kim, S. W. *et al.* Identification of functionally important amino acid residues within the C2 domain of human factor V using alanine scanning mutagenesis. *Biochemistry* (in the press).
12. Ortel, T. L. *et al.* Localization of functionally important epitopes within the second C-type domain of coagulation factor V using recombinant chimeras. *J. Biol. Chem.* **269**, 15898–15905 (1994).
13. Ortel, T. L. *et al.* Inhibitory anti-factor V antibodies bind to the factor V C2 domain and are associated with hemorrhagic manifestations. *Blood* **91**, 4188–4196 (1998).
14. Rosing, J., Bakker, H. M., Thomassen, M. C., Hemker, H. C. & Tans, G. Characterization of two forms of human factor Va with different cofactor activities. *J. Biol. Chem.* **268**, 21130–21136 (1993).
15. Koppaka, V., Talbot, W. F., Zhai, X. & Lentz, B. R. Roles of factor Va heavy and light chains in protein and lipid rearrangements associated with the formation of a bovine factor Va-membrane complex. *Biophys. J.* **73**, 2638–2652 (1997).
16. Ito, N., Phillips, S. E. V., Yadav, K. D. S. & Knowles, P. F. Crystal structure of a free radical enzyme, galactose oxidase. *J. Mol. Biol.* **238**, 794–814 (1994).
17. Baumgartner, S., Hofmann, K., Chiquet-Ehrismann, R. & Bucher, P. The discoidin domain family revisited—new members from prokaryotes and a homology-based fold prediction. *Protein Sci.* **7**, 1626–1631 (1998).
18. Krishnaswamy, S. & Mann, K. G. The binding of factor Va to phospholipid vesicles. *J. Biol. Chem.* **263**, 5714–5723 (1988).
19. Swairjo, M. A., Concha, N. O., Kaetzel, M. A., Dedman, J. R. & Seaton, B. A. Ca²⁺-bridging mechanism and phospholipid head group recognition in the membrane-binding protein annexin V. *Nature Struct. Biol.* **2**, 968–974 (1995).
20. Gerads, L., Govers-Riemsag, J. W. P., Tans, G., Zwaal, R. F. A. & Rosing, J. Prothrombin activation on membranes with anionic lipids containing phosphate, sulfate, and/or carboxyl groups. *J. Biol. Chem.* **29**, 7967–7974 (1990).
21. Rosing, J., Speijer, H. & Zwaal, R. F. A. Prothrombin activation on phospholipid membranes with positive electrostatic potential. *Biochemistry* **27**, 8–11 (1988).
22. Cutsforth, G. A. *et al.* Insights into the complex association of bovine factor Va with acidic-lipid-containing synthetic membranes. *Biophys. J.* **70**, 2938–2949 (1996).
23. Kalafatis, M., Rand, M. D. & Mann, K. G. Factor Va—membrane interaction is mediated by two regions located on the light chain of the cofactor. *Biochemistry* **33**, 486–493 (1994).
24. Gilbert, G. E., Furie, B. C. & Furie, B. Binding of human factor VIII to phospholipid vesicles. *J. Biol. Chem.* **265**, 815–822 (1990).
25. Leslie, A. in *Crystallographic Computing V* (eds Moras, D., Podjarny, A. D. & Thierry, J. C.) 27–38 (Oxford Univ. Press, 1991).
26. Collaborative Computational Project No. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
27. Knight, S. Ribulose 1,5-bisphosphate carboxylase/oxygenase—a structural study. (PhD Thesis, Swedish Univ. Agricultural Sciences, Uppsala, 1989).
28. Brünger, A. T. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1993).
29. Navaza, J. An automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
30. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

Acknowledgements

We thank D. Grosse and M. Renatus for initial crystallization trials. W.B. is supported by the Biomed. and Training and Mobility Programs of the EU, by the HFSP and by the SFB469; M.A.Q.-A. and W.H.K. by the NIH; T.L.O., a Pew Scholar, by the March of Dimes Birth Defects Foundation; and S.M.-R. by PraxisXXI, FCT, Portugal.

Correspondence and requests for materials should be addressed to W.B. (e-mail: bode@biochem.mpg.de). Atomic coordinates have been deposited in the Protein Data Bank under accession codes 1CZS, 1CZT and 1CZV.

Structure of the C2 domain of human factor VIII at 1.5 Å resolution

Kathleen P. Pratt^{*,†}, Betty W. Shen^{*}, Kazuya Takeshima[†], Earl W. Davie[†], Kazuo Fujikawa[†] & Barry L. Stoddard^{*}

^{*} Program in Structural Biology, Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N. A3-023, Seattle, Washington 98109, USA

[†] Department of Biochemistry, University of Washington, Box 357350, Seattle, Washington 98195, USA

Human factor VIII is a plasma glycoprotein that has a critical role in blood coagulation^{1,2}. Factor VIII circulates as a complex with von Willebrand factor³. After cleavage by thrombin, factor VIIIa associates with factor IXa at the surface of activated platelets or endothelial cells^{4,5}. This complex activates factor X (refs 6, 7), which in turn converts prothrombin to thrombin in the presence of factor Va (refs 8, 9). The carboxyl-terminal C2 domain of factor VIII contains sites that are essential for its binding to von Willebrand factor and to negatively charged phospholipid surfaces^{4,10,11}. Here we report the structure of human factor VIII C2 domain at 1.5 Å resolution. The structure reveals a β -sandwich core, from which two β -turns and a loop display a group of solvent-exposed hydrophobic residues. Behind the hydrophobic surface lies a ring of positively charged residues. This motif suggests a mechanism for membrane binding involving both hydrophobic and electrostatic interactions. The structure explains, in part, mutations in the C2 region of factor VIII that lead to bleeding disorders in haemophilia A.

The secondary structure and overall fold of the C2 domain of factor VIII are shown in Fig. 1. The domain contains 19 β -strands, 8 of which form a β -sandwich core fold. The structure of the β -sandwich is similar to the lipid-binding domain of galactose oxidase and to the homology model based on that structure^{12,13} (the r.m.s. difference between the two superimposed β -sandwich cores of the factor VIII C2 domain and galactose oxidase domain is 1.8 Å). Three β hairpins and an additional loop extend beyond the β -sandwich fold. These elements of the structure were not predicted by homology modelling; in general, the structure of the protein outside the β -sandwich core differs considerably from the galactose oxidase lipid-binding domain. The C2 domain displays one turn of 3₁₀ helix near its amino terminus. The N- and C-terminal regions are linked by a disulphide bridge between residues Cys 2,174 and Cys 2,326, and a *cis* peptide bond precedes Pro 2,299. In addition to the homology to the galactose oxidase lipid-binding domain, the factor VIII C2 structure also resembles regions within neuraminidase (Protein Data Bank code 1eut), with an r.m.s. difference of 1.8 Å. Despite their similar nomenclature, the C2 domains of factors V and VIII are distinct from the 'C2 domains' of signal-transduction proteins such as protein kinase C and phospholipases C and A2.

The protein structure exhibits two extensive regions of exposed hydrophobic surface (Fig. 2). The first region, at the upper end of the β -sandwich, may represent a surface that interacts with the C1 domain in the factor VIIIa molecule, or with other components of the factor IXa enzyme complex. The second hydrophobic surface is formed primarily by the two β -hairpins at the opposite end of the molecule. The first hairpin presents Met 2,199 and Phe 2,200 to the solvent. The second hairpin turn places Leu 2,551 and Leu 2,252 at the same surface. Val 2,223 extends from an adjacent loop and also contributes to this surface. Underlying this hydrophobic region is a ring of at least four basic residues. We propose that docking of the C2 domain on the membrane surface entails burial of hydrophobic residues in the membrane and the interaction of basic residues with negatively charged phospholipid head groups.

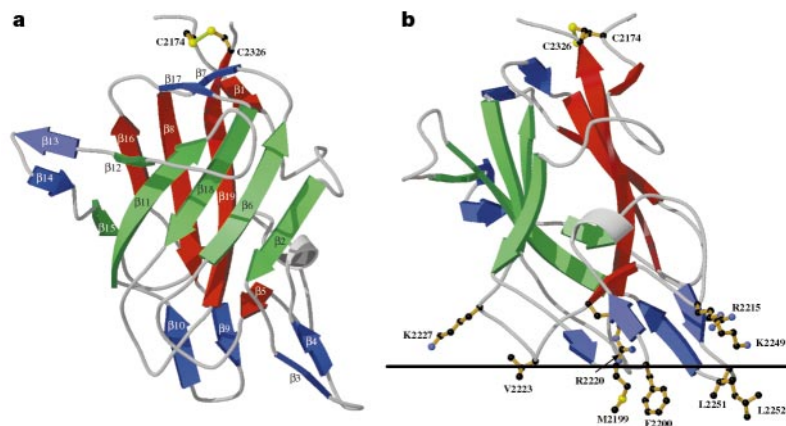


Figure 1 Ribbon diagram and proposed membrane-binding mode. **a**, The structure consists of 19 β -strands, including an eight-stranded β -sandwich core. Two of the three β -hairpins and an additional loop extend beyond the core fold. These regions, which are predominantly hydrophobic, flank a pair of positively charged clefts. **b**, The putative

Previous studies reported that a 21-residue peptide corresponding to residues 2,303–2,323 of the C2 domain competes with factor VIII for membrane binding^{14,15}. NMR measurements demonstrated that this peptide assumes an amphipathic helical structure in detergent micelles, leading to the hypothesis that a similar structure might be formed by these residues in the protein–membrane complex. However, the crystal structure reveals that these residues constitute a part of the β -sandwich core. Conversion of this region to an α -helix would involve a major conformational rearrangement.

Twenty-one residues in the factor VIII C2 domain are known to be sites of deleterious individual point mutations in patients with haemophilia A (ref. 16) (Fig. 3). Notably, only one of these mutation sites, Val 2,223, is located in the region that we suggest is directly involved in membrane binding. We propose that the design of the membrane-binding region is highly redundant, with multiple hydrophobic and basic side chains contributing to the binding energy¹⁷. The mutation of a single solvent-exposed side chain in this region would generally result in a reduced, but still effective, binding affinity to phospholipid bilayers. The most important contributions to the membrane-binding free energy would come from solvation

membrane-binding surface includes M2199 and F2200 from the first turn, L2251 and L2252 from the second turn, and V2223 from the adjacent loop. These hydrophobic residues and a belt of basic residues partition on either side of the horizontal grey line which indicates the putative polar/nonpolar boundary of a phospholipid bilayer surface.

changes and electrostatic-charge neutralization, neither of which are strongly dependent on the precise orientation of particular amino-acid side chains.

In contrast, mutations in the protein core or at protein surfaces that interact with other regions of factor VIII or von Willebrand factor could interfere with protein folding, transport or secretion. This might lead to an increase in degradation and clearance from the serum, and result in more profound physiological consequences. Of the reported C2 point mutations in haemophiliacs, six are at

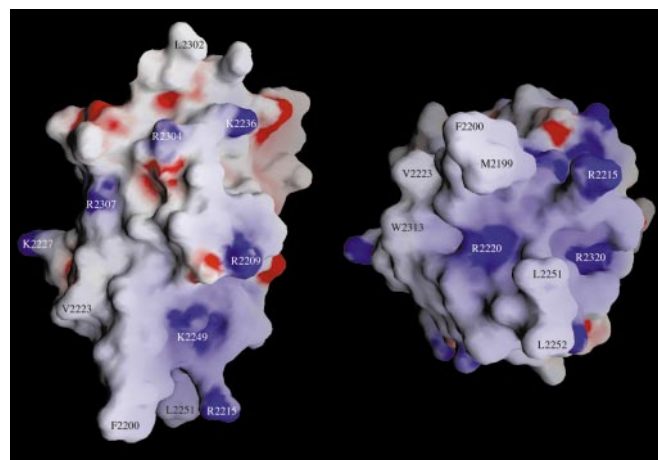


Figure 2 Molecular surface of the factor VIII C2 domain. Left panel, orientation similar to that in Fig. 1a; right panel, the structure is rotated by $\sim 90^\circ$ to look directly into the bottom of the molecule. Blue represents positively charged residues; red, negatively charged; white, uncharged. Note the uncharged protrusions formed by the turns and loops shown in Fig. 1. W2313 also appears to participate in this hydrophobic surface. A 'ring' of solvent-accessible, positively charged residues lies directly behind this surface. The hydrophobic patch on the top of the molecule is also seen in the left panel.

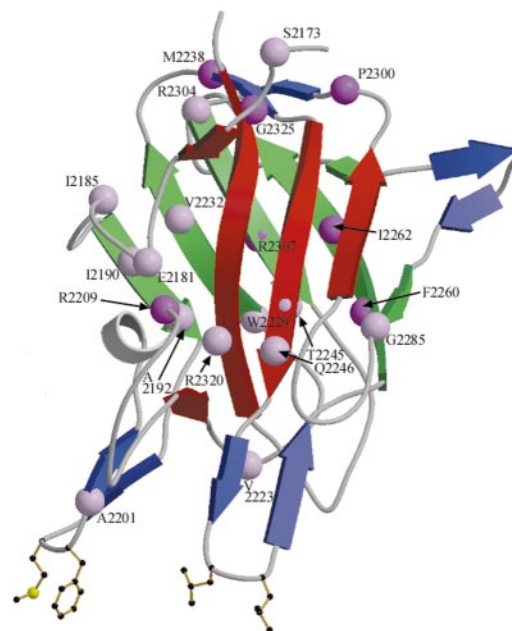


Figure 3 Location of haemophilia-A-associated point mutations in the factor VIII C2 domain. The mutation sites correspond to: (1) residues forming the protein core (I2,185, I2,190, V2,232, T2,245, F2,260 and I2,262); (2) positions at the second hydrophobic surface (P2,300, M2,238); (3) other surface-exposed residues (S2,173, E2,181, A2,192, A2,201, V2,223, W2,229, G2,285, R2,304 and R2,307); and (4) other residues that appear to be structurally important (R2,209, Q2,246, R2,320 and G2,325). Sites of mutations that have been associated with severe haemophilia A are shown as purple spheres, while sites associated with mild to moderate bleeding disorders are pink. Four sites shown in this figure (S2,173, A2,201, V2,232 and R2,320) have not been previously described (A. Thompson, personal communication). Clinical symptoms of individuals carrying the same mutation can vary significantly.

Table 1 Data and refinement statistics

	Native 1	Hg	Pt	Native 2
Diffraction data				
Resolution (Å)	2.2	2.2	1.7	1.5
Source	RAXIS-IV	RAXIS-IV	ALS 5.0.2	NSLS X-26C
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell <i>a</i>	45.9	45.8	46.2	46.4
<i>b</i>	57.0	57.0	57.2	56.4
<i>c</i>	66.2	66.3	66.2	65.7
Wavelength (Å)	1.54	1.54	1.07	1.01
No. of refl. (unique)	7,938	8,134	19,711	27,232
Redundancy	5.3	4.8	6.1	4.5
Completeness*	97.0 (94.6)	99.6 (97.3)	91.2 (83.4)	96.0 (94.0)
R_{merge} (%)	4.2	5.6	3.6	4.0
MIR phasing				
No. of sites		1	1	
$R_{\text{iso}}, K_{\text{amp}}$		18.2, 6.6	18.0, 11.5	
Phasing power†		1.2/1.3/0.9	0.9/1.0/1.9	
R_{cullis}		0.8/0.8/0.9	0.8/0.9/0.6	
Overall FOM‡			0.36/0.31	
Refinement				
Resolution range				50.0–1.5
R_{cryst}				20.5
R_{free}				21.5
No. of protein atoms				1253
No. of solvent atoms				240
Ramachandran distribution§				86.0, 12.0,
(% core, allowed, generous, disallowed)				1.3, 0.7
R.m.s. bonds, angles				0.005 Å, 1.47°
Average (B)				17.7 Å ²

* Completeness reported for all reflection and for the highest-resolution shell.

† Phasing power and R_{cullis} reported for centric isomorphous differences, acentric isomorphous differences and acentric anomalous differences, respectively.

‡ Overall figure of merit reported for acentric and centric reflections, respectively.

§ One residue (His 2,315) is in the disallowed region of the Ramachandran plot; this residue is in density of excellent quality and its geometry is supported in all unbiased density maps.

residues that comprise the hydrophobic protein core. With the exception of Ile 2,185, point mutations at these sites are associated with reduced factor VIII antigen levels and/or with moderate to severe bleeding defects. Four additional mutation sites, at residues Arg 2,209, Gln 2,246, Arg 2,320 and Gly 2,325 affect residues that are involved in hydrogen-bonding networks that clearly stabilize the protein fold. Met 2,238 and Pro 2,300 are located at the second hydrophobic surface. Finally, three exposed side chains (Trp 2,229, Arg 2,304 and Arg 2,307) are at a common surface distal to both the putative membrane-association region and the other hydrophobic interface. Unlike the residues buried in the protein core, these exposed residues do not appear critical for protein folding. Point mutations at all three of these positions are associated with bleeding disorders. Possibly, this surface represents a binding site for another protein, such as von Willebrand factor. Mutations that interfere with the complex between factor VIII and von Willebrand factor would be expected to result in phenotypes similar to those caused by structurally destabilizing mutations, because the clearance rate for free factor VIII is considerably higher than that of the factor VIII in association with von Willebrand factor. Alternatively, this surface on the C2 domain may interact with other regions of factor VIIIa or factor IXa.

Many haemophilia A patients who receive repeated factor VIII infusions develop auto-antibodies and allo-antibodies that inhibit the activity of factor VIII, making the replacement therapy less effective¹⁸. A major determinant of this inhibitor response has been mapped to the C2 domain^{19–21}. The C2 inhibitors interfere with the binding of factor VIIIa to phospholipid and to von Willebrand factor²². The identification of surface-exposed amino acids should now assist in the identification of specific residues that make up this epitope. □

Methods

The structure was solved by multiple isomorphous replacement (MIR), using a secreted protein construct (residues 2,171–2,332) that incorporates a cysteine-to-valine substitui-

tion at its N terminus and a point mutation, S2296C, that was introduced for the purpose of mercury derivatization. The resulting construct was cloned in frame with the α -factor secretion signal sequence, to express the recombinant proteins under the control of the AOX1 promoter, which drives secretion in *Pichia pastoris* strain GS115 (Invitrogen, Inc.). The protein was purified according to the protocol for the wild-type C2 domain¹¹.

Crystals of the recombinant S2296C-C2 protein were grown from 2 M ammonium sulphate, 0.1 M MES pH 6.0, 0.015% Na₂S₂O₈. The crystals were frozen after transfer to a cryoprotectant buffer of similar composition that contained 30% glucose w/v and 10% glycerol v/v. The crystals displayed considerable non-isomorphism despite reproducible unit-cell dimensions. Therefore, native and isomorphous derivative data were collected from single crystals that were subjected to sequential rounds of cryocooling, data collection, thawing and metal soaks. Two derivatives were prepared for the S2296C-C2 protein mutant: the first by soaking a crystal in cryobuffer containing 2 mM of the monovalent mercurial reagent *p*-hydroxymercuriphenylsulphonic acid (PSMB, Sigma); the second by soaking another crystal in a 1 mM K₂PtCl₄ solution.

The native and mercury derivative data were collected to 2.2 Å resolution on an in-house RAXIS-IV area detector, whereas the platinum derivative data were collected to 1.7 Å, corresponding to the platinum anomalous edge. An additional native data set was collected to 1.5 Å resolution at Brookhaven NSLS beamline X-25 and used for the refinement. All data were processed using DENZO/SCALEPACK²³ and merged using the CCP4 program suite²⁴; phases were calculated and refined using SHARP²⁵, SOLOMON²⁴ and 2D histogram matching²⁶. Model building was done with O (ref. 27), and the structure was refined using CNS²⁸ after removing 5% of the measurements to monitor the free *R*-factor²⁹. The final R_{cryst} is 20.5% and the R_{free} is 21.5%. The refined model of the protein domain consists of 159 amino acids and 240 water molecules. After model building and refinement, it was discovered that the mercurial reagent PSMB was bound near the internal disulphide bond. The stereochemical quality of the protein model was examined throughout the refinement using PROCHECK³⁰. Data and refinement statistics are shown in Table 1.

Received 26 August; accepted 27 October 1999.

- Kane, W. H. & Davie, E. W. Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders. *Blood* **71**, 539–555 (1988).
- Kaufman, R. J. Biological regulation of factor VIII activity. *Annu. Rev. Med.* **43**, 325–339 (1992).
- Fay, P. J., Coumans, J.-V. & Walker, F. J. Von Willebrand factor mediates protection of factor VIII from activated protein C-catalyzed inactivation. *J. Biol. Chem.* **266**, 2172–2177 (1991).
- Foster, P. A., Fulcher, C. A., Houghten, R. A. & Zimmerman, T. S. Synthetic factor VIII peptides with amino acid sequences contained within the C2 domain of factor VIII inhibit factor VIII binding to phosphatidylserine. *Blood* **75**, 1999–2004 (1990).
- Duffy, E. J., Parker, E. T., Mutucumarana, V. P., Johnson, A. E. & Lollar, P. Binding of factor VIIIa and factor VIII to factor IXa on phospholipid vesicles. *J. Biol. Chem.* **267**, 17006–17011 (1992).
- Duffy, E. J. & Lollar, P. Intrinsic pathway activation of factor X and its activation peptide-deficient derivative, factor Xdes-143-191. *J. Biol. Chem.* **267**, 7821–7827 (1992).
- van Diejen, G., Tans, G., Rosing, J. & Hemker, H. C. The role of phospholipid and factor VIIIa in the activation of bovine factor X. *J. Biol. Chem.* **266**, 3433–3442 (1991).
- Krishnaswamy, S., Nesheim, M. E., Prydzial, E. L. G. & Mann, K. G. Assembly of prothrombinase complex. *Methods Enzymol.* **222**, 260–280 (1993).
- Mann, K. G., Nesheim, M. E., Church, W. R., Haley, P. & Krishnaswamy, S. Surface-dependent reactions of the vitamin K-dependent enzyme complexes. *Blood* **76**, 1–16 (1990).
- Saenko, E. L. & Scandella, D. A mechanism for inhibition of factor VIII binding to phospholipid by von Willebrand factor. *J. Biol. Chem.* **270**, 13826–13833 (1995).
- Takeshima, K. & Fujikawa, K. The phospholipid binding property of the C2 domain of human factor VIII. *J. Biol. Chem.* (submitted).
- Pellequer, J.-L., Gale, A. J., Griffin, J. H. & Getzoff, E. D. Homology models of the C domains of blood coagulation factors V and VIII: a proposed membrane binding mode for FV and FVIII C domains. *Blood Cells Mol. Dis.* **24**, 448–461 (1998).
- Baumgartner, S., Hofmann, K., Chiquet-Ehrismann, R. & Bucher, P. The discoidin domain family revisited: new members from prokaryotes and a homology-based fold prediction. *Protein Sci.* **7**, 1626–1631 (1998).
- Veeraghavan, S., Baleja, J. D. & Gilbert, G. E. Structure and topography of the membrane-binding C2 domain of factor VIII in the presence of dodecylphosphocholine micelles. *Biochem. J.* **332**, 549–555 (1998).
- Gilbert, G. E. & Baleja, J. D. Membrane-binding peptide from the C2 domain of factor VIII forms an amphipathic structure as determined by NMR spectroscopy. *Biochemistry* **34**, 3022–3031 (1995).
- Haemostasis Research Group at the MRC Clinical Sciences Centre, Imperial College Medical School. Haemophilia A mutation database. <http://europium.mrc.rpms.ac.uk/HAMsterS>.
- Wolfenden, R., Andersson, L., Cullis, P. M. & Southgate, C. C. B. Affinities of amino acid side chains for solvent water. *Biochemistry* **20**, 849–855 (1981).
- Hoyer, L. W. & Scandella, D. Factor VIII inhibitors: structure and function in autoantibody and hemophilia A patients. *Semin. Hematol.* **31** (Suppl. 4), 1–5 (1994).
- Healey, J. F. et al. Residues Glu2181–Val2243 contain a major determinant of the inhibitory epitope in the C2 domain of human factor VIII. *Blood* **92**, 3701–3709 (1998).
- Jacquemin, M. G. et al. Mechanism and kinetics of factor VIII inactivation: study with an IgG4 monoclonal antibody derived from a hemophilia A patient with inhibitor. *Blood* **92**, 496–506 (1998).
- Kuwabara, I. et al. Mapping of the minimal domain encoding a conformational epitope by lambda phage surface display: factor VIII inhibitor antibodies from haemophilia A patients. *J. Immun. Methods* **224**, 89–99 (1999).
- Shima, M. et al. A factor VIII neutralizing monoclonal antibody and a human inhibitor alloantibody recognizing epitopes in the C2 domain inhibit factor VIII binding to von Willebrand factor and to phosphatidylserine. *Thromb. Haemost.* **69**, 240–246 (1993).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computing Project 4. The CCP4 Suite: Programs for protein crystallography. *Acta Cryst. D* **50**, 760–763 (1994).

25. LaFortelle, E., Irwin, J. J. & Bricogne, G. in *Crystallographic Computing* Vol. 7 (eds Bourne, P. & Watenpaugh, K.) 100–130 (Oxford Science Publications, Oxford, 1997).
26. Nieh, Y.-P. & Zhang, K. Y. J. A 2D histogram matching method for protein phase refinement and extension. *Acta Crystallogr. D* (in the press).
27. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
28. Brunger, A. T. *et al.* Crystallography and NMR System. *Acta Crystallogr. D* **44**, 905–921 (1998).
29. Brunger, A. Assessment of phase accuracy by cross validation: the free R value. Methods and Applications. *Acta Crystallogr. D* **49**, 24–36 (1993).
30. Laskowski, R. J., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–290 (1993).

Acknowledgements

We thank J. Bolduc, R. Strong, Y.-P. Nieh, K. Zhang, L. Steward, H. Kim and A. Thompson for advice and assistance. This work was funded by the NIH.

Correspondence and requests for materials should be addressed to B.L.S. (e-mail: bstoddar@fhcrc.org). The coordinates of the structure have been deposited in the Protein Data Bank (accession code = 1D7P).

addendum

Energetic constraints on the diet of terrestrial carnivores

Chris Carbone, Georgina M. Mace, S. Craig Roberts & David W. Macdonald

Nature **402**, 286–288 (1999)

S.C.R.'s present address is Evolution and Behaviour Research Group, Department of Psychology, University of Newcastle, Newcastle NE1 7RU, UK. □

Please release me

This software update includes releases of old favourites like SigmaPlot.

ChemOffice 2000

CambridgeSoft www.camsoft.com

A suite of chemistry tools including ChemDraw 5, Chem3D and ChemFinder

Available in Ultra and Pro versions, ChemOffice 2000 is a PC-based desktop working environment for chemists, combining tools for drawing, modelling and information handling. The Ultra 2000 version offers new features including the ability to estimate NMR spectra and convert name-to-structure (Name = Struct) with ChemDraw 5.

Reader Service No. 100

Origin 6.0

Microcal www.microcal.com

Bilingual data analysis and technical graphics software

Two updates this year take Microcal Software's Windows-based Origin package to version 6.0. Joining the Origin 6.0 is German Origin 6.0. This provides all the Origin 6.0 and Origin 6.0 Professional functions, with the addition of a completely translated German Tutorial and User Manual. Features common to both packages include data point masking tools for controlled analysis, a new interface for customizing toolbars, and a comprehensive Plot Details dialogue box for easy customization of all graphical elements.

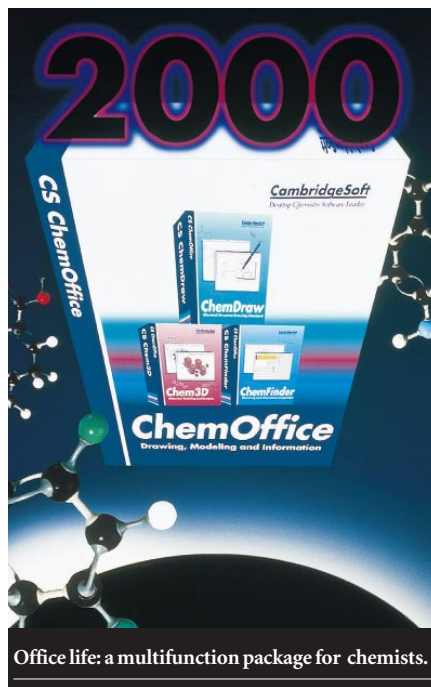
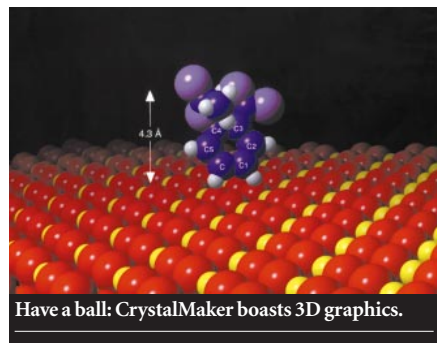
Reader Service No. 101

CrystalMaker

Crystal Software www.crystallmaker.co.uk

A major upgrade to this crystal and molecular structures visualization program

CrystalMaker version 4.0 boasts enhanced 3D graphics: wearing the supplied stereo glasses, crystals and molecules "appear to leap out of the screen and hover in mid air", and they can be rotated and manipulated in real time. New tools permit manipulation and design of new structures or modelling of defects. Other new features include display of vectors, smoothed graphics, flexibility for



displaying and editing individual atoms and bonds, and support for new Mac system technologies including extended file dialogues (Navigation Services) with graphical previews and the MacOS 'platinum' appearance.

Reader Service No. 102

GelExpert 3.5

NucleoTech Corporation

www.nucleotech.com

Latest incarnation of an imaging package

Three new analysis modules are added in release 3.5 of Gel Expert, Protein Gel Analysis, Genotyping and BioAssay. GelExpert, a package to capture, process and analyse images, is now available for Windows 95, 98 and NT. The software has extensive file format compatibility. Images can be acquired from virtually any video camera, from Twain compatible scanners and from a variety of high-sensitivity cooled CCD cameras. Features include an image library, multiple plug-in analysis modules for 1D, 2D, QPCR, microbial colony counting, and densitometric analysis.

Reader Service No. 103

Metrodata TiNet 2.3

Metrohm

www.metrohm.ch

Software for computer-supported titration

TiNet 2.3 is a Windows program (for 95, 98 or NT) that supports all aspects of titration: method development, sample data memory, real-time titration curve, documentation and saving of results; reprocessing of the mea-

sured data and exporting of results to a LIMS.

Reader Service No. 104

MagNet

Infolytica

www.infolytica.com

A longest-established electromagnetics simulation package, now in version 6.0

This new PC release features a completely rewritten interface to exploit the native properties of Windows. This means that MagNet v6 can now exchange data with other programs to perform optimization. Using Microsoft's ActiveX technology, scripts may be developed for automatic model building, typically using Visual Basic Script language or Java.

Reader Service No. 105

IML Tools

Windmill Software

www.windmill.co.uk

Save time creating measurement, control and analysis applications

IML Tools can be used directly from Windows applications like Excel or Access, or from javascript code running on a web page. They also speed up programming in Visual Basic, with functions to configure the instrument, read channels and write to channels. The Tools are 32-bit and run under Windows 2000, NT, 98 or 95. They are bundled with ready-to-run applications from the Windmill software suite for logging, charting, output control and dynamic data exchange. IML Tools communicate directly to instruments from Bruel & Kjaer, Mettler Toledo, Sartorius and other manufacturers.

Reader Service No. 106

Drug File

Derwent Information

www.derwent.co.uk

Online data on pharmaceuticals

The Derwent Drug File, operating through Ovid software, contains over 1.6 million references dating back to 1964 from pharmaceuticals journals. The biological and chemical drug research information available in this system covers all aspects of drug development, synthesis, evaluation, manufacture and use. Each record includes an abstract outlining research methods and detailing results, plus systematic indexing.

Reader Service No. 107

Vector NTI

Informax

www.informaxinc.com

Software for sequencing and the like

The latest version of Vector NTI Suite 5.5 molecular biology software, released in

August, can be sampled as a free demo via the company's website. Both Mac and PC versions are available. The suite includes BioPlot for protein and nucleic-acid sequence analysis, AlignX for multiple sequence analysis and Vector NTI for data management, mapping, primer design and other functions.

Reader Service No. 108

Nuclides 2000

Institute for Transuranium Elements

http://itumagill.fzk.de/nuclides_2000

An electronic nuclide chart on CD-ROM

This Windows-compatible CD contains the 'Nuclide Explorer and Decay Engine', with information on over 2,600 radionuclides accessed through a periodic table/Segrè chart. Also on the CD is a multimedia library covering aspects of radionuclides from radio-carbon dating to supernovae, with related websites hotlinked.

Reader Service No. 109

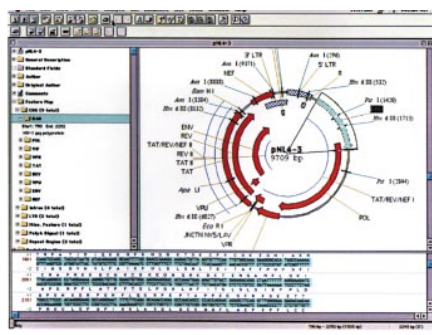
gNMR 4.1

Cherwell Scientific

www.cherwell.com

Latest version of a program to analyse spectral data

Version 4.1 of gNMR for Windows, the NMR simulation software package, is designed for chemists who perform NMR analysis and



Vector 5.5: organize those sequences.

simulation and who require fast results on their PC. Database functionality allows the user to store data in Paradox/dbase and ODBC database formats. Reference deconvolution makes it possible to create more attractive spectra, with modest line narrowing. Maximum system size is increased from 23 to 49 nuclei. An online demo is available via Cherwell's website.

Reader Service No. 110

SigmaPlot 5.0

SPSS Science

www.spss.com

The graphics program hits version 5.0

SPSS are now shipping SigmaPlot 5.0 for Windows. This version offers increased automation to allow the application to be customized for a researcher's specific needs. The graphics enable users to create exact 2D and 3D technical graphs for presenting research results. The most powerful new tool is the macro recorder that allows macros to be recorded with a simple point and click. Other enhanced features include: OLE automation to give access to SigmaPlot seamlessly from stand-alone applications like Word and Excel.

Reader Service No. 111

STATISTICA'99

StatSoft Ltd

www.statsoft.com

Customizable statistics software

The STATISTICA'99 Statistical Solutions package combines a large array of general as well as specialized statistical and graphical methods. The latest version includes html support, enhanced connectivity options, implementation of Visual General Linear Models, Visual Generalised Linear Models, Visual General Stepwise Regression and Visual Partial Least Squares.

Reader Service No. 112

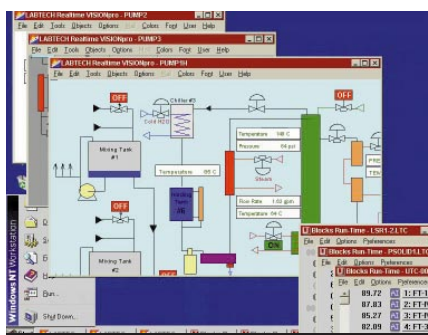
NOTEBOOK/CONTROL

Labtech

www.labtech.com

New Windows versions of Labtech's instrumentation and control software

NOTEBOOK, NOTEBOOKpro, CONTROL and CONTROLpro are now available for Windows NT/Windows 2000. New capabilities include true multiprocessing, priority-



In CONTROL: upgrades from Labtech.

based scheduling, and a new open architecture design. Multiprocessing allows multiple copies of NOTEBOOK and CONTROL to run on the same system. The software is connected by hotlinks to data sources through LT-Speedway, the company's high-speed software bus.

Reader Service No. 113

Procite 5 for Windows

ISI ResearchSoft

www.isiresearchsoft.com

Upgrade for an well-established bibliographical tool

This major upgrade to ProCite turns the bibliographic program into an online research tool, allowing users to search remote bibliographic databases on the Internet. ProCite 5 adds an important new feature to the information toolbox, the ability to access reference data on the Internet and easily incorporate it into personal databases and documents. The new program ships with more than 200 Internet Host files that instantly link the user to databases such as university card catalogues, the Library of Congress, PubMed, PsycINFO, and ERIC from providers such as Ovid Technologies, SilverPlatter, and OCLC.

Reader Service No. 114

CpGWare

Intergen

www.intergenco.com/cpgware

Software for DNA methylation analysis

CpGWare Primer Design Software simplifies the design of primers for methylation specific PCR (MSP), a powerful technique for investigating the methylation status of gene promoters. CpGWare is designed to be used online. The program output consists of primer sequence sets capable of discriminating between methylated and un-methylated DNA, template DNA sequence with the primer locations identified, and relevant amplicon sizes. Subscriptions giving unlimited access for 6 or 12 months are available, and a free demo is accessible via Intergen's website.

Reader Service No. 115

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

ADVERTISEMENTS

HIGH QUALITY PEPTIDE SYNTHESIS

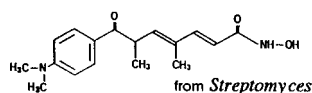
- 'Standard' and 'non-standard'
- Fast delivery
- Competitive prices
- Also advisory services

Pepsyn Ltd

Contact: Dr J A Smith, Chairman, Pepsyn Ltd, School of Biological Sciences, Life Science Building, The University of Liverpool L69 7ZB, UK
Tel: 0151 794 4329 Fax: 0151 794 4349
e-mail: J.A.Smith@liverpool.ac.uk
Website: www.pepsyn.co.uk

READER ENQUIRY NO. 221

Trichostatin A



Catalog No. 204-11991

This histone deacetylase inhibitor is highly effective at low concentrations, with high specificity. For more information, please call
Wako BioProducts: (800) 992-9256; (804) 271-7677.

READER ENQUIRY NO. 218